

**Appl. No.** : 10/063,517  
**Filed** : May 7, 2002

### **REMARKS**

Applicants have submitted herewith a substitute specification in compliance with the requirements set forth in 37 C.F.R. §1.58(c) and in M.P.E.P. § 6808.01. The substitute specification contains no new matter.

Claims 1-5 remain present for examination. Applicants respond below to the specific rejections raised by the Examiner in the final Office Action mailed March 31, 2005.

#### **Specification:**

The Examiner has not entered the substitute specification submitted on January 18, 2005. According to the Examiner, Applicant failed to provide a statement that the substitute specification includes no new matter, and failed to provide a marked up copy showing the changes relative to the prior version. The Examiner noted that the specification omitted the claim for priority. The Examiner has also objected to the specification on the grounds that it contains embedded hyperlinks.

Applicants submit herewith a substitute specification. The specification includes no new matter. As indicated by the appropriate markings, all hyperlinks have been removed from the specification, and all trademarks are represented in capital letters. The specification includes the paragraph in which Applicants assert their claim to priority. Applicants submit that the substitute specification is in compliance with 37 C.F.R. § 1.125, and respectfully request that it be entered accordingly.

#### **Rejection under 35 U.S.C. §101 – Utility**

The Examiner has rejected Claims 1-5 as allegedly not being supported by a specific, substantial and credible, or well-established utility for the reasons set forth in the Office Action mailed October 4, 2004. The Examiner asserts that the data provided in the specification do not support the conclusion that the gene encoding PRO300 or the PRO300 polypeptide is differentially expressed in lung cancer compared to normal lung tissue such that the PRO300 polypeptides or PRO300 antibodies could be used as a diagnostic tool. The Examiner also states that since it is not known whether the nucleic acid is involved in tumor suppression or if it is inhibited by tumor progression, one skilled in the art could not use the PRO300 nucleic acids, PRO300 polypeptides, or PRO300 antibodies as therapeutic or diagnostic tools. Further, the

**Appl. No.** : 10/063,517  
**Filed** : May 7, 2002

Examiner states that absent information regarding what type of lung tumors were analyzed in Example 18, the data are preliminary and do not confer utility on the claimed invention.

The Examiner maintains that the two Grimaldi declarations, the Polakis declaration, and the Ashkenazi declaration are insufficient to overcome the rejections set forth by the Examiner in the October 4, 2004 Office Action.

The Examiner argues that the first Grimaldi declaration is unpersuasive because one skilled in the art would not conclude that a gene whose cDNA levels differ by at least two-fold in lung tumor versus normal lung tissue would be useful as a diagnostic marker for lung cancer. First, according to the Examiner, an increase in DNA copy number does not necessarily correlate with a cancerous state. In support of this line of reasoning, the Examiner cites to Hittleman *et al.* (2001), Ann. N.Y. Acad. Sci. 952:1-12, for the proposition that “because aneuploid DNA can be found in normal tissue, detection of increased DNA copy number does not necessarily mean that those cells containing the DNA are cancerous.” Office Action at 6. According to the Examiner, it is not clear from the data presented in Example 18 whether “PRO300 is amplified in cancerous lung epithelium more than in damaged (non cancerous) lung epithelium.” *Id.* Further, the Examiner argues that not all lung tumor tissue would be expected to have the same expression pattern for a given gene, and since the specification fails to indicate what types of lung tumor tissue were used in the Example 18, the data are preliminary and insufficient to establish utility.

The Examiner also rejects the testimony set forth in the second Grimaldi declaration regarding the correlation between gene overexpression and protein overexpression, arguing that “these statements have no basis in fact and appear to be the opinion of the Declarant.” Office Action at 7. In addition, the Examiner asserts that “there has been no distinction on the record in general or in the specification as filed between total nucleic acid, which includes chromosomal DNA and mRNA.” *Id.* at 8. The Examiner contends that one cannot determine from the data presented whether “the observed ‘amplification’ of nucleic acid is due to increase in chromosomal copy number, or alternatively due to an increase in transcription rates.” *Id.* Citing Pennica *et al.*, the Examiner states that increased DNA copy numbers are not necessarily correlated with increased mRNA levels.

Additionally, the Examiner makes the distinction between Dr. Grimaldi’s discussion of the relationship between chromosomal aberrations, mRNA levels and protein overexpression for Her2/Neu, and the claimed invention, noting that unlike Her2/Neu, PRO300 has not been

**Appl. No.** : 10/063,517  
**Filed** : May 7, 2002

associated with tumor formation or cancer, and that no translocation of the gene encoding PR0300 is known to occur. The Examiner concludes that the data presented in Example 18 are preliminary. The Examiner also rejects Dr. Grimaldi's statement that even if mRNA and protein levels do not correlate, then measurement of gene and protein expression levels enables more accurate tumor classification, "because only the cDNA was measured, because of aneuploidy, because no information is provided on expression levels of the protein, because there is no information on what type of lung tumor tissues were tested, etc." Office Action at 12.

The Examiner maintains that the Polakis declaration is unpersuasive. The Examiner first asserts that "it is important to note that the instant specification provides no information regarding increased mRNA levels of PRO300 in tumor samples relevant [sic] to normal samples. . .[t]herefore the declaration is insufficient to overcome the rejection of the claims." Office Action at 9-10. The Examiner then refers to Hu *et al.* (2003), J. Proteome Res. 2:405-412, for the proposition that there is "no evidence of a correlation between [genes differentially expressed in cancer versus normal tissue] and a known role in the disease." Office Action at 10. The Examiner also argues that for genes with low mRNA levels, the mRNA levels are not predictive of protein levels. Citing to Gygi *et al.* the Examiner argues that "message levels are generally not predictive of protein levels for the majority of genes studies [sic] by Gygi *et al.*" *Id.* at 11.

Finally, the Examiner argues that the Ashkenazi declaration is insufficient to overcome the utility rejection. Specifically, the Examiner rejects Dr. Ashkenazi's testimony that simultaneous testing of gene expression and gene product levels enables more accurate tumor classification "because it has not been demonstrated that the protein of the instant invention is differentially expressed in different tumors. If it was, the protein would have a specific and substantial utility." Office Action at 13. The Examiner then states that "[o]verexpression of a gene product in a cancer cell does not necessarily mean that the gene product is involved in the cancer and that targeting the gene product would also be therapeutic." *Id.* at 14. The Examiner reiterates that the data in Example 18 do not provide information on levels of protein expression, and concludes that thus the proposed use is simply a starting point for further research.

Applicants respectfully disagree.

Appl. No. : 10/063,517  
Filed : May 7, 2002

Utility – Legal Standard

According to the Utility Examination Guidelines (“Utility Guidelines”), 66 Fed. Reg. 1092 (2001) an invention complies with the utility requirement of 35 U.S.C. § 101, if it has at least one asserted “specific, substantial, and credible utility” or a “well-established utility.”

Under the Utility Guidelines, a utility is “specific” when it is particular to the subject matter claimed. For example, it is generally not enough to state that a nucleic acid is useful as a diagnostic tool without also identifying the condition that is to be diagnosed.

The requirement of “substantial utility” defines a “real world” use, and derives from the Supreme Court’s holding in *Brenner v. Manson*, 383 U.S. 519, 534 (1966) stating that “The basic *quid pro quo* contemplated by the Constitution and the Congress for granting a patent monopoly is the benefit derived by the public from an invention with substantial utility.” In explaining the “substantial utility” standard, M.P.E.P. § 2107.01 cautions, however, that Office personnel must be careful not to interpret the phrase “immediate benefit to the public” or similar formulations used in certain court decisions to mean that products or services based on the claimed invention must be “currently available” to the public in order to satisfy the utility requirement. “Rather, *any reasonable use that an applicant has identified for the invention that can be viewed as providing a public benefit should be accepted as sufficient*, at least with regard to defining a ‘substantial’ utility.” (M.P.E.P. § 2107.01, emphasis added).

The mere consideration that further experimentation might be performed to more fully develop the claimed subject matter does not support a finding of lack of utility. M.P.E.P. § 2107.01 III cites *In re Brana*, 51 F.3d 1560, 1566, 34 U.S.P.Q.2d 1436 (Fed. Cir. 1995) in stating that “Usefulness in patent law ... necessarily includes the expectation of further research and development. The stage at which an invention in this field becomes useful is well before it is ready to be administered to humans.” Further, “[T]o violate § 101 the claimed device must be totally incapable of achieving a useful result” *Juicy Whip Inc. v. Orange Bang Inc.*, 51 USPQ2d 1700 (Fed. Cir. 1999), citing *Brooktree Corp. v. Advanced Micro Devices, Inc.*, 977 F.2d 1555, 1571 (Fed.Cir.1992).

Indeed, the Guidelines for Examination of Applications for Compliance With the Utility Requirement, set forth in M.P.E.P. § 2107 II(B)(1) gives the following instruction to patent examiners: “If the applicant has asserted that the claimed invention is useful for any particular



Appl. No. : 10/063,517  
Filed : May 7, 2002

practical purpose ... and the assertion would be considered credible by a person of ordinary skill in the art, do not impose a rejection based on lack of utility.”

Utility need NOT be Proved to a Statistical Certainty – a Reasonable Correlation between the Evidence and the Asserted Utility is Sufficient

An Applicant's assertion of utility creates a presumption of utility that will be sufficient to satisfy the utility requirement of 35 U.S.C. § 101, “unless there is a reason for one skilled in the art to question the objective truth of the statement of utility or its scope.” *In re Langer*, 503 F.2d 1380, 1391, 183 USPQ 288, 297 (CCPA 1974). *See, also In re Jolles*, 628 F.2d 1322, 206 USPQ 885 (CCPA 1980); *In re Irons*, 340 F.2d 974, 144 USPQ 351 (1965); *In re Sichert*, 566 F.2d 1154, 1159, 196 USPQ 209, 212-13 (CCPA 1977). Compliance with 35 U.S.C. § 101 is a question of fact. *Raytheon v. Roper*, 724 F.2d 951, 956, 220 USPQ 592, 596 (Fed. Cir. 1983) cert. denied, 469 US 835 (1984). The evidentiary standard to be used throughout *ex parte* examination in setting forth a rejection is a preponderance of the evidence, or “more likely than not” standard. *In re Oetiker*, 977 F.2d 1443, 1445, 24 USPQ2d 1443, 1444 (Fed. Cir. 1992). This is stated explicitly in the M.P.E.P.:

[T]he applicant does not have to provide evidence sufficient to establish that an asserted utility is true “beyond a reasonable doubt.” **Nor must the applicant provide evidence such that it establishes an asserted utility as a matter of statistical certainty.** Instead, evidence will be sufficient if, considered as a whole, it leads a person of ordinary skill in the art to conclude that the asserted utility is more likely than not. M.P.E.P. at § 2107.02, part VII (2004) (underline emphasis in original, bold emphasis added, internal citations omitted).

The PTO has the initial burden to offer evidence “that one of ordinary skill in the art would reasonably doubt the asserted utility.” *In re Brana*, 51 F.3d 1560, 1566, 34 U.S.P.Q.2d 1436 (Fed. Cir. 1995). Only then does the burden shift to the Applicant to provide rebuttal evidence. *Id.* As stated in the M.P.E.P., such rebuttal evidence does not need to absolutely prove that the asserted utility is real. Rather, the evidence only needs to be reasonably indicative of the asserted utility.

In *Fujikawa v. Wattanasin*, 93 F.3d 1559, 39 U.S.P.Q.2d 1895 (Fed. Cir. 1996), the Court of Appeals for the Federal Circuit upheld a PTO decision that *in vitro* testing of a novel pharmaceutical compound was sufficient to establish practical utility, stating the following rule:

[T]esting is often required to establish practical utility. But the test results **need not absolutely prove** that the compound is pharmacologically active. All that is

Appl. No. : 10/063,517  
Filed : May 7, 2002

required is that the tests be “*reasonably* indicative of the desired [pharmacological] response.” In other words, there must be a **sufficient correlation** between the tests and an asserted pharmacological activity so as to convince those skilled in the art, **to a reasonable probability**, that the novel compound will exhibit the asserted pharmacological behavior.” *Fujikawa v. Wattanasin*, 93 F.3d 1559, 1564, 39 U.S.P.Q.2d 1895 (Fed. Cir. 1996) (internal citations omitted, bold emphasis added, italics in original).

While the *Fujikawa* case was in the context of utility for pharmaceutical compounds, the principals stated by the Court are applicable in the instant case where the asserted utility is for a therapeutic and diagnostic use – utility does not have to be established to an absolute certainty, rather, the evidence must convince a person of skill in the art “to a reasonable probability.” In addition, the evidence need not be direct, so long as there is a “sufficient correlation” between the tests performed and the asserted utility.

The Court in *Fujikawa* relied in part on its decision in *Cross v. Iizuka*, 753 F.2d 1040, 224 U.S.P.Q. 739 (Fed. Cir. 1985). In *Cross*, the Appellant argued that basic *in vitro* tests conducted in cellular fractions did not establish a practical utility for the claimed compounds. Appellant argued that more sophisticated *in vitro* tests using intact cells, or *in vivo* tests, were necessary to establish a practical utility. The Court in *Cross* rejected this argument, instead favoring the argument of the Appellee:

[I]n *vitro* results...are generally predictive of *in vivo* test results, i.e., there is a **reasonable correlation** therebetween. Were this not so, the testing procedures of the pharmaceutical industry would not be as they are. [Appellee] has not urged, and rightly so, that there is an invariable exact correlation between *in vitro* test results and *in vivo* test results. Rather, [Appellee's] position is that successful *in vitro* testing for a particular pharmacological activity establishes a **significant probability** that *in vivo* testing for this particular pharmacological activity will be successful. *Cross v. Iizuka*, 753 F.2d 1040, 1050, 224 U.S.P.Q. 739 (Fed. Cir. 1985) (emphasis added).

The *Cross* case is very similar to the present case. Like *in vitro* testing in the pharmaceutical industry, those of skill in the field of biotechnology rely on the reasonable correlation that exists between gene expression and protein expression (see below). Were there no reasonable correlation between the two, the techniques that measure gene levels such as microarray analysis, differential display, and quantitative PCR would not be so widely used by those in the art. As in *Cross*, Applicants here do not argue that there is “an invariable exact correlation” between gene expression and protein expression. Instead, Applicants’ position

Appl. No. : 10/063,517  
Filed : May 7, 2002

detailed below is that a measured change in gene expression in cancer cells establishes a “significant probability” that the expression of the encoded polypeptide in cancer will also be changed based on “a reasonable correlation therebetween.”

Taken together, the legal standard for demonstrating utility is a relatively low hurdle. An Applicant need only provide evidence such that it is **more likely than not that a person of skill in the art would be convinced, to a reasonable probability, that the asserted utility is true.** The evidence need not be direct evidence, so long as there is a reasonable correlation between the evidence and the asserted utility. The Applicant **does not need to provide evidence such that it establishes an asserted utility as a matter of statistical certainty.**

Finally, in assessing the credibility of the asserted utility, the M.P.E.P. states that “to overcome the presumption of truth that an assertion of utility by the applicant enjoys” the PTO must establish that it is “more likely than not that one of ordinary skill in the art would doubt (i.e., “question”) the truth of the statement of utility.” M.P.E.P. § 2107.02 III A. The M.P.E.P. cautions that:

Rejections under 35 U.S.C. 101 have been **rarely sustained** by federal courts. Generally speaking, **in these rare cases**, the 35 U.S.C. 101 rejection was sustained either because the **applicant ... asserted a utility that could only be true if it violated a scientific principle, such as the second law of thermodynamics, or a law of nature, or was wholly inconsistent with contemporary knowledge in the art.** M.P.E.P. § 2107.02 III B. (underline emphasis in original, bold emphasis added); citing *In re Gazave*, 379 F.2d 973, 978, 154 USPQ 92, 96 (CCPA 1967).

Even assuming that the PTO has met its initial burden to offer evidence that one of ordinary skill in the art would reasonably doubt the truth of the asserted utility, Applicants assert that they have met their burden of providing rebuttal evidence such that it is more likely than not those skilled in the art, to a reasonable probability, would believe that the claimed invention is useful as a diagnostic tool for cancer.

*The Data in Example 18 are Data Regarding Differential mRNA Levels, not Gene Amplification*

Applicants begin by clarifying that the data concerning the differential expression of the PRO300 gene presented in Example 18 relate to gene expression, **not gene amplification**. The description of Example 18 makes clear that the results were obtained by quantitative PCR amplification of cDNA libraries. It is well known in the art that cDNA libraries are made from

Appl. No. : 10/063,517  
Filed : May 7, 2002

mRNA, and reflect the level of mRNA for a particular gene in the source tissue. Thus, Example 18 is reporting a measure of the *expression* of the PRO300 gene, i.e. mRNA levels, not its *amplification*, i.e. the number of copies of PRO300 in the genome.

Throughout the final Office Action, the Examiner refers to “gene amplification.” For example, the Examiner states “[b]ecause aneuploid DNA can be found in normal tissue, detection of increased DNA copy number does not necessary [sic] mean those cells containing the DNA are cancerous. The gene amplification assay disclosed in the instant specification does not provide a comparison between the lung tumor samples and normal lung epithelium control,” (Office Action at 6), “all that the specification does is present evidence that the DNA encoding PRO300 is amplified,” (Office Action at 7-8), “[since] there has been no distinction on the record in general or in the specification as filed between total nucleic acid, which includes chromosomal DNA, and mRNA. . .one cannot determine from the data in the specification whether the observed ‘amplification’ of nucleic acid is due to increase in chromosomal copy number, or alternatively due to an increase in transcription rates,” (Office Action at 8), and “it is important to note that the instant specification provides no information regarding increased mRNA levels of PRO300 in tumor samples relevant [sic] to normal samples. Only gene amplification data was presented in Example 18.” Office Action at 9-10. In rejecting Applicant’s previous arguments and evidence, the Examiner asserts, “it does not necessarily follow that an increase in gene copy number results in increased gene expression and increased polypeptide expression.” *Id.* at 8.

Applicants submit that the Examiner has misinterpreted that data presented in this application. Clearly, Example 18 reports data regarding differential mRNA levels, **not** gene amplification data.

Applicants wish to clarify that whether or not gene amplification leads to increased gene expression is irrelevant to this particular application. Applicants have provided reliable evidence that the PRO300 mRNA is differentially expressed in certain tumors. Whether this differential expression is due to changes in gene copy number, transcription rates, a combination of the two, or some other known or unknown cellular mechanism is simply not relevant to Applicants’ asserted utility.

#### Summary of Applicants’ Arguments and the PTO’s Response

In an attempt to clarify Applicants’ argument, Applicants offer a summary of their argument and the disputed issues involved. Applicants assert that the claimed antibodies have

**Appl. No.** : 10/063,517  
**Filed** : May 7, 2002

utility as diagnostic tools for cancer, particularly lung cancer. Applicants are not asserting that the claimed antibodies necessarily provide a definitive diagnosis of lung cancer, but rather that they are useful, alone or in combination with other diagnostic tools to assist in the diagnosis of lung cancers. Applicants' asserted utility rests on the following argument:

1. Applicants have provided reliable evidence that mRNA for the PRO300 polypeptide is expressed at least two-fold higher in normal lung tissue compared to lung tumor.
2. Applicants assert that it is well-established in the art that a change in the level of mRNA for a particular protein, e.g. a decrease, generally leads to a corresponding change in the level of the encoded protein, e.g. a decrease;
3. Given Applicants' evidence that the level of mRNA for the PRO300 polypeptide is decreased in lung tumor compared to normal lung tissue, it is likely that the level of PRO300 polypeptide is also decreased in lung tumors and therefore antibodies to PRO300 are useful as a diagnostic tool to distinguish lung tumor from normal lung tissue.

Applicants understand the Examiner to be making several arguments in response to Applicants' asserted utility:

1. The Examiner argues that the data provided in Example 18 does not support the conclusion that the gene encoding PRO300 or PRO300 polypeptides are differentially expressed in lung cancer compared to normal tissue such that the PRO300 polynucleotides, polypeptides, and antibodies are useful as a diagnostic or therapeutic tool;
2. The Examiner argues that different types of lung tumor tissue would not be expected to have the same gene expression pattern, and since the specification does not indicate what types of lung tumor tissue were used in Example 18, the data are insufficient to establish utility;
3. The Examiner argues that because no mutation or translocation of PRO300 has been associated with lung cancer, and because it is not known whether the nucleic acid is involved in tumor suppression or is inhibited by tumor progression, the disclosure is insufficient to meet the utility requirement;
4. The Examiner cites Hittleman *et al.* to support her assertion that because aneuploid DNA can be found in normal tissue, detection of increased DNA copy number does not necessarily correlate with a cancerous state, and relies on Pennica *et al.* as further evidence of the lack of correlation between gene amplification and protein levels;

Appl. No. : 10/063,517  
Filed : May 7, 2002

5. The Examiner cites Hu *et al.* and Gygi *et al.* to support her assertion that the literature cautions against drawing conclusions based on small changes in transcription expression levels between normal and cancerous tissue, and that mRNA levels are not predictive of protein levels.

As detailed below, Applicants submit that the PTO has failed to meet its initial burden to offer evidence “that one of ordinary skill in the art would reasonably doubt the asserted utility.” *In re Brana*, 51 F.3d 1560, 1566, 34 U.S.P.Q.2d 1436 (Fed. Cir. 1995). First, the Examiner has failed to offer relevant evidence to support her rejection of the data in Example 18 and the Declaration of Chris Grimaldi in support of these data. Second, the fact that no mutation or translocation of PRO300 has been associated with lung cancers does not prevent PRO300 or PRO300 antibodies from being useful as a diagnostic tool for the same. Third, Applicants submit that given the well-established correlation between a change in the level of mRNA with a corresponding change in the levels of the encoded protein, the PRO300 protein is likely differentially expressed in lung tumors. This provides utility for PRO300 and related proteins and PRO300 antibodies as cancer diagnostic tools. Finally, even if the PTO has met its initial burden, Applicants have submitted enough rebuttal evidence to establish that it is **more likely than not** that a person of skill in the art would be convinced, **to a reasonable probability**, that the asserted utility is true. As stated above, Applicants’ evidence need not be direct evidence, so long as there is a reasonable correlation between the evidence and the asserted utility. **The standard is not statistical or absolute certainty.**

*Applicants have established that the Gene Encoding the PRO300 Polypeptide is Differentially Expressed in Certain Cancers compared to Normal Tissue*

Applicants first address the PTO’s argument that the evidence of differential expression of the gene encoding the PRO300 polypeptide lung tumors is insufficient.

The gene expression data in the specification, Example 18, shows that the mRNA associated with protein PRO300 was more highly expressed in normal lung tissue compared to lung tumor. Gene expression was analyzed using standard semi-quantitative PCR amplification reactions of cDNA libraries isolated from different human tumor and normal human tissue samples. Identification of the differential expression of the PRO300 polypeptide-encoding gene in lung tumor tissue compared to the corresponding normal lung tissue renders the molecule useful as a diagnostic tool for the determination of the presence or absence of lung tumor. In support, Applicants previously submitted as Exhibit 1 a first Declaration of J. Christopher

Appl. No. : 10/063,517  
Filed : May 7, 2002

Grimaldi, an expert in the field of cancer biology. This declaration explains the importance of the data in Example 18, and how differential gene and protein expression studies are used to differentiate between normal and tumor tissue (see Declaration, paragraph 7).

In paragraph 5 of his declaration, Mr. Grimaldi states that the gene expression studies reported in Example 18 of the instant application were made from pooled samples of normal and of tumor tissues. Mr. Grimaldi explains that:

The DNA libraries used in the gene expression studies were made from pooled samples of normal and of tumor tissues. *Data from pooled samples is more likely to be accurate than data obtained from a sample from a single individual.* That is, the detection of variations in gene expression is likely to represent a more generally relevant condition when pooled samples from normal tissues are compared with pooled samples from tumors in the same tissue type. (Paragraph 5) (emphasis added).

The Examiner argues that “different types of cancer demonstrate different gene expression patterns and the measurement of a single gene would not necessarily be indicative of the presence or absence of cancer.” Applicants disagree, and submit that the Examiner is essentially requiring Applicants to show data that the polypeptides can be used to detect or differentiate different types of lung cancer. Such a showing is not necessary for the asserted utilities of the claimed antibodies, e.g., diagnostic tools to distinguish between tumor tissue and normal tissue. Applicants assert that the differential expression data in Example 18 provide evidence that decreased expression of the PRO300 gene can provide a useful diagnostic tool for lung cancer, i.e. that the PRO300 gene “can be used to differentiate tumor from normal” (Grimaldi declaration, paragraph 7), and maintain that the Examiner’s assertion that Applicants have not described what types of lung tumors were assayed in Example 18 does not detract from Applicants’ asserted utility.

In paragraphs 6 and 7, Mr. Grimaldi explains that the semi-quantitative analysis employed to generate the data of Example 18 is sufficient to determine if a gene is over- or under-expressed in tumor cells compared to corresponding normal tissue. He states that any visually detectable difference seen between two samples is indicative of at least a two-fold difference in cDNA between the tumor tissue and the counterpart normal tissue. Thus, the results of Example 18 reflect at least a two-fold difference between normal and tumor samples. He also states that the results of the gene expression studies indicate that the genes of interest “can be used to differentiate tumor from normal,” thus establishing their reliability. He explains

**Appl. No.** : 10/063,517  
**Filed** : May 7, 2002

that, contrary to the Examiner's assertions, "The precise levels of gene expression are irrelevant; what matters is that there is a relative difference in expression between normal tissue and tumor tissue." (Paragraph 7). Thus, since it is the relative level of expression between normal tissue and suspected cancerous tissue that is important, the precise level of expression in normal tissue is irrelevant. Likewise, there is no need for quantitative data to compare the level of expression in normal and tumor tissue. As Mr. Grimaldi states, "If a difference is detected, this indicates that the gene and its corresponding polypeptide and antibodies against the polypeptide are useful for diagnostic purposes, to screen samples to differentiate between normal and tumor."

Applicants submit that the declaration of Mr. Grimaldi is based on personal knowledge of the relevant facts at issue. Mr. Grimaldi is an expert in the field and conducted or supervised the experiments at issue. Applicants remind the Examiner that "[o]ffice personnel must accept an opinion from a qualified expert that is based upon relevant facts whose accuracy is not being questioned." PTO Utility Examination Guidelines (2001) (emphasis added). In addition, declarations relating to issues of fact should not be summarily dismissed as "opinions" without an adequate explanation of how the declaration fails to rebut the Examiner's position. *In re Alton* 76 F.3d 1168 (Fed. Cir. 1996). The Examiner has not supplied any reasons or evidence to question the accuracy of the facts upon which Mr. Grimaldi based his opinion, other than the Examiner's assertions that aneuploid DNA can be found in normal tissue, and that not all tumor tissue would be expected to have the same expression pattern for a given gene. See, Office Action at 6-7. As discussed above, evidence relating to DNA copy number is irrelevant to the issue at hand, as is an assertion that Applicants have not detailed the types of lung tumors tested. Mr. Grimaldi has personal knowledge of the relevant facts, has based his opinion on those facts, and the Examiner has offered no reason or evidence to reject either the underlying facts or his opinion. Applicants submit that for the reasons stated above, the Examiner's arguments are unpersuasive, and the Examiner should accept Mr. Grimaldi's opinion with regard to his statement that "any visually detectable difference seen between two samples is indicative of at least a two-fold difference in cDNA between the tumor tissue and the counterpart normal tissue" and that the genes of interest "can be used to differentiate tumor from normal." Together, these statements establish that there is at least a two-fold difference in expression, and that the results are reliable enough that they can be used to distinguish tumor from normal tissue. Finally, Applicants submit that whether this differential expression is due to an increase in chromosomal



**Appl. No.** : 10/063,517  
**Filed** : May 7, 2002

copy number or increased transcription rates is not relevant to whether the difference in expression can be used to distinguish tumor from normal tissue.

In conclusion, Applicants submit that the evidence reported in Example 18, combined with the first Grimaldi Declaration, establish that there is at least a two-fold difference in PRO300 cDNA between lung tumor tissue and normal lung tissue. Therefore, it follows that expression levels of the PRO300 gene can be used to distinguish lung tumor tissue from normal lung tissue. The Examiner has not offered any significant arguments or evidence to the contrary.

As Applicants explain below, it is more likely than not that the PRO300 polypeptide is also differentially expressed in lung tumor, and can therefore be used to distinguish the same tumor tissue from normal lung counterparts. This provides utility for the claimed PRO300 antibodies.

*Applicants have established that the Accepted Understanding in the Art is that there is a Direct Correlation between mRNA Levels and the Level of Expression of the Encoded Protein*

Applicants next turn to the second portion of their argument in support of their asserted utility – that it is well-established in the art that a change in the level of mRNA for a particular protein generally leads to a corresponding change in the level of the encoded protein; given Applicants’ evidence of differential expression of the mRNA for the PRO300 polypeptide in lung, it is more likely than not that the PRO300 polypeptide is differentially expressed; and proteins differentially expressed in lung tumor have utility as diagnostic tools.

In support of the assertion that changes in mRNA are positively correlated to changes in protein levels, Applicants previously submitted a copy of a second Declaration by J. Christopher Grimaldi, an expert in the field of cancer biology (previously attached as Exhibit 2). As stated in paragraph 5 of the declaration, “Those who work in this field are well aware that in the vast majority of cases, when a gene is over-expressed...the gene product or polypeptide will also be over-expressed.... This same principal applies to gene under-expression.” Further, “the detection of increased mRNA expression is expected to result in increased polypeptide expression, and the detection of decreased mRNA expression is expected to result in decreased polypeptide expression. The detection of increased or decreased polypeptide expression can be used for cancer diagnosis and treatment.” The references cited in the declaration and submitted herewith support this statement.

Appl. No. : 10/063,517  
Filed : May 7, 2002

Applicants also previously submitted a copy of the declaration of Paul Polakis, Ph.D. (previously attached as Exhibit 3), an expert in the field of cancer biology. As stated in paragraph 6 of his declaration:

Based on my own experience accumulated in more than 20 years of research, including the data discussed in paragraphs 4 and 5 above [showing a positive correlation between mRNA levels and encoded protein levels in the vast majority of cases] and my knowledge of the relevant scientific literature, it is my considered scientific opinion that for human genes, an increased level of mRNA in a tumor cell relative to a normal cell typically correlates to a similar increase in abundance of the encoded protein in the tumor cell relative to the normal cell. In fact, *it remains a central dogma in molecular biology that increased mRNA levels are predictive of corresponding increased levels of the encoded protein.* (Emphasis added).

Dr. Polakis acknowledges that there are published cases where such a correlation does not exist, but states that it is his opinion, based on over 20 years of scientific research, that “such reports are exceptions to the commonly understood general rule that increased mRNA levels are predictive of corresponding increased levels of the encoded protein.” (Polakis Declaration, paragraph 6).

The statements of Grimaldi and Polakis are supported by the teachings in Molecular Biology of the Cell, a leading textbook in the field (Bruce Alberts, *et al.*, Molecular Biology of the Cell (3<sup>rd</sup> ed. 1994) (submitted herewith as Exhibit 1) and (4<sup>th</sup> ed. 2002) (previously submitted as Exhibit 4, hereinafter “Alberts, 4<sup>th</sup> Ed.”). Figure 9-2 of Exhibit 1 shows the steps at which eukaryotic gene expression can be controlled. The first step depicted is transcriptional control. Exhibit 1 provides that “[f]or most genes transcriptional controls are paramount. This makes sense because, of all the possible control points illustrated in Figure 9-2, only transcriptional control ensures that no superfluous intermediates are synthesized.” Exhibit 1 at 403 (emphasis added). In addition, the text states that “Although controls on the initiation of gene transcription are the predominant form of regulation for most genes, other controls can act later in the pathway from RNA to protein to modulate the amount of gene product that is made.” Exhibit 1 at 453 (emphasis added). Thus, as established in Exhibit 1, the predominant mechanism for regulating the amount of protein produced is by regulating transcription initiation.

In Alberts, 4<sup>th</sup> Ed., Figure 6-3 on page 302 illustrates the basic principle that there is a correlation between increased gene expression and increased protein expression. The accompanying text states that “a cell can change (or regulate) the expression of each of its genes

Appl. No. : 10/063,517  
Filed : May 7, 2002

according to the needs of the moment – *most obviously by controlling the production of its mRNA.*” Alberts, 4<sup>th</sup> Ed. at 302 (emphasis added). Similarly, Figure 6-90 on page 364 of Alberts, 4<sup>th</sup> Ed. illustrates the path from gene to protein. The accompanying text states that while potentially each step can be regulated by the cell, “the initiation of transcription is the most common point for a cell to regulate the expression of each of its genes.” Alberts, 4<sup>th</sup> Ed. at 364 (emphasis added). This point is repeated on page 379, where the authors state that of all the possible points for regulating protein expression, “[f]or most genes transcriptional controls are paramount.” Alberts, 4<sup>th</sup> Ed. at 379 (emphasis added).

Further support for Applicants’ position can be found in the textbook, *Genes VI*, (Benjamin Lewin, *Genes VI* (1997)) (submitted herewith as Exhibit 2) which states “having acknowledged that control of gene expression can occur at multiple stages, and that production of RNA cannot inevitably be equated with production of protein, it is clear that the overwhelming majority of regulatory events occur at the initiation of transcription.” *Genes VI* at 847-848 (emphasis added).

Additional support is also found in Zhigang *et al.*, *World Journal of Surgical Oncology* 2:13, 2004, submitted herewith as Exhibit 3. Zhigang studied the expression of prostate stem cell antigen (PSCA) protein and mRNA to validate it as a potential molecular target for diagnosis and treatment of human prostate cancer. The data showed “a high degree of correlation between PSCA protein and mRNA expression” Exhibit 3 at 6. Of the samples tested, 81 out of 87 showed a high degree of correlation between mRNA expression and protein expression. The authors conclude that “it is demonstrated that PSCA protein and mRNA overexpressed in human prostate cancer, and that the increased protein level of PSCA was resulted from the upregulated transcription of its mRNA.” Exhibit 3 at 11. Even though the correlation between mRNA expression and protein expression occurred in 93% of the samples tested, not 100%, the authors state that “PSCA may be a promising molecular marker for the clinical prognosis of human Pca and a valuable target for diagnosis and therapy of this tumor.” *Id.*

Further, Meric *et al.*, *Molecular Cancer Therapeutics*, vol. 1, 971-979 (2002), submitted herewith as Exhibit 4, states the following:

The **fundamental principle** of molecular therapeutics in cancer is to exploit the differences in gene expression between cancer cells and normal cells...[M]ost efforts have concentrated on identifying differences in gene expression at the level

**Appl. No.** : 10/063,517  
**Filed** : May 7, 2002

of mRNA, which can be attributable to either DNA amplification or to differences in transcription. Meric *et al.* at 971 (emphasis added).

Those of skill in the art would not be focusing on differences in gene expression between cancer cells and normal cells if there were no correlation between gene expression and protein expression.

Together, the declarations of Grimaldi and Polakis, the accompanying references, and the excerpts and references provided above all establish that the accepted understanding in the art is that there is a reasonable correlation between changes in gene expression and the level of the encoded protein.

In response to the second Grimaldi Declaration and the Polakis Declaration, the Examiner makes three arguments. First, the Examiner states that unlike Her2/Neu and t(5;14) discussed in the second Grimaldi declaration, the PRO300 gene has not been associated with tumor formation, and no translocation of PRO300 is known to occur. The Examiner concludes that in the absence of any of the above information, the disclosure is insufficient to meet the utility requirement. Second, the Examiner maintains that “one cannot determine from the data in the specification whether the observed ‘amplification’ of nucleic acid is due to increase in chromosomal copy number, or alternatively due to an increase in transcription rates.” Office Action at 8. Finally, the Examiner cites to Pennica, Hu, and Gygi in support of her position that increased DNA copy number does not necessarily correlate with increased gene product level, that there is no correlation between differential mRNA expression levels and a known role in cancer, and that transcript levels provide little predictive value with respect to levels of protein expression. Applicants address each argument in turn.

Applicants submit that a lack of known role for PRO300 in cancer does not prevent its use as a diagnostic tool for cancer. The fact that there is no known translocation or mutation of PRO300 is irrelevant to whether its differential expression can be used to assist in diagnosis of cancer – one does not need to know why PRO300 is differentially expressed, or what the consequence of the differential expression is, in order to exploit the differential expression to distinguish tumor from normal tissue. In fact the Revised Interim Utility Guidelines promulgated by the PTO recognize that proteins which are differentially expressed in cancer have utility. (See the caveat in Example 12 which state that the utility requirement is satisfied where a protein is expressed in melanoma cells but not on normal skin and antibodies against the protein can be

**Appl. No.** : 10/063,517  
**Filed** : May 7, 2002

used to diagnose cancer.) In light of the above, Applicants maintain that the Examiner's position that "over-expression of a gene product in a cancer cell does not necessarily mean that the gene product itself is involved in the cancer and that targeting the gene product could also be therapeutic," is contrary to the PTO guidelines. In addition, while Applicants appreciate that actions taken in other applications are not binding on the PTO with respect to the present application, Applicants note that the PTO has issued several patents claiming differentially expressed polypeptides. (*See, e.g.*, U.S. Patent No. 6,414,117 and U.S. Patent No. 6,124,433, attached hereto as Exhibits 5 and 6.)

In response to the Examiner's second argument, Applicants reiterate that the data in Example 18 are gene expression data, not gene amplification data. The specification and the first Grimaldi Declaration make clear that Example 18 used semi-quantitative PCR of cDNA libraries, and disclose differential PRO300 gene expression, not gene amplification. As discussed above, the Examiner cites Pennica *et al.* for support of the argument that gene amplification does not necessarily lead to increased gene expression. This reference is irrelevant to the instant application which reports differential gene expression, not gene amplification. Therefore, the reference and the Examiner's arguments regarding gene amplification do not support the Examiner's challenge of the sufficiency of the Example 18 data, or the first Grimaldi Declaration. Likewise, the Examiner's reliance on Hittleman *et al.* is equally ineffective to cast any doubt on Applicants' asserted utilities and evidence in support thereof. As with Pennica *et al.*, Hittleman *et al.* relates to aneuploidy and DNA copy number, and is simply irrelevant to the gene expression data presented in Example 18. As such, neither Pennica *et al.* or Hittleman *et al.* cast doubt on Applicants' asserted utilities.

In further support of her rejection of Applicants' asserted utilities and testimony, the Examiner cites Hu *et al.* (2003, *J. Proteome Research*, 2: 405-412) for the proposition that "the literature cautions researchers from drawing conclusions based on small changes in transcript expression levels between normal and cancerous tissue." Office Action at 10. Applicants submit that Hu is uninformative regarding the correlation between gene expression level and protein levels. In Hu, the researchers used an automated literature-mining tool to summarize and estimate the relative strengths of all human gene-disease relationships published on Medline. They then generated a microarray expression dataset comparing breast cancer and normal breast tissue. Using their data-mining tool, they looked for a correlation between the strength of the

**Appl. No.** : 10/063,517  
**Filed** : May 7, 2002

literature association between the gene and breast cancer, and the magnitude of the difference in expression level. They report that for genes displaying a 5-fold change or less in tumors compared to normal, there was no evidence of a correlation between altered gene expression and a *known* role in the disease. See Hu at 411. However, among genes with a 10-fold or more change in expression level, there was a strong correlation between expression level and a *published* role in the disease. *Id.* at 412. Importantly, Hu reports that the observed correlation was only found among estrogen receptor-positive tumors, not ER-negative tumors. *Id.*

The general findings of Hu are not surprising – one would expect that genes that have the greatest change in expression in a disease would be the first targets of research, and therefore have the strongest *known* relationship to the disease as measured by the number of reports of a connection in the literature. But this does not mean that genes, and their corresponding proteins, with a lower level of change in expression are not important or cannot be used as molecular markers of the disease. This is demonstrated by the fact that ER-negative tumors did not show a correlation. The correlation reported in Hu only indicates that the greater the change in expression level, the more likely it is that there is a *published* or *known* role for the gene in the disease, as found by their automated literature-mining software. Nowhere in Hu does it say that a lack of correlation in their study means that the genes, and their corresponding proteins, with a less than five-fold change in level of expression in cancer cannot serve as a molecular marker of cancer. Genes with lower levels of change in expression may or may not be the most important genes in causing the disease, but the genes and their corresponding proteins can still show a consistent and measurable change in expression. While such genes and polypeptides may or may not be good targets for further research, they can nonetheless be used as diagnostic tools. Thus, Hu does not refute the Applicants' assertion that the PRO300 gene and its corresponding polypeptide can be used as a cancer diagnostic tool because they are differentially expressed in melanoma relative to normal skin tissue.

Finally, the Examiner cites Gygi *et al.* for the proposition that “transcript levels provide little predictive value with respect to the extent of protein expression,” and thus one skilled in the art would doubt the principle that mRNA levels and protein levels correlate. Office Action at 11. Although Gygi states that “there was a general trend of increased protein levels resulting from increased mRNA levels,” (Gygi, p. 1726), the Examiner argues that the correlation observed in Gygi “was highly biased by a small number of genes with very large protein and message levels.”

Appl. No. : 10/063,517  
Filed : May 7, 2002

Office Action at 11. Applicants note that Gygi states that the high degree of variability seen at low levels of mRNA is due to the fact that "the magnitude of the error in the measurement of mRNA levels is inversely proportional to the mRNA levels." Gygi, p. 1727. Applicants also note that Gygi states that the correlation between mRNA and protein levels is especially strong for highly expressed mRNAs. Gygi, p. 1726. Considering that Example 18 of the specification shows over-expression of PRO300 mRNA in normal lung, Gygi actually provides strong evidence in support of a general correlation between mRNA and protein levels, and thus the utility of the claimed PRO300 antibodies.

In view of the above, Applicants submit that the references cited by the Examiner do not cast doubt on Applicants' asserted utilities or the testimony of Dr. Grimaldi and Dr. Polakis. To the contrary, the cited references support Applicants' position that more likely than not, one skilled in the art would not doubt Applicants' asserted utility.

*The Arguments made by the Examiner are Not Sufficient to satisfy the PTO's Initial Burden of Offering Evidence "that one of ordinary skill in the art would reasonably doubt the asserted utility"*

As stated above, an Applicant's assertion of utility creates a presumption of utility that will be sufficient to satisfy the utility requirement of 35 U.S.C. § 101, "unless there is a reason for one skilled in the art to question the objective truth of the statement of utility or its scope." *In re Langer*, 503 F.2d 1380, 1391, 183 USPQ 288, 297 (CCPA 1974). The evidentiary standard to be used throughout *ex parte* examination in setting forth a rejection is a preponderance of the evidence, or "more likely than not" standard. *In re Oetiker*, 977 F.2d 1443, 1445, 24 USPQ2d 1443, 1444 (Fed. Cir. 1992). This is stated explicitly in the M.P.E.P.:

[T]he applicant does not have to provide evidence sufficient to establish that an asserted utility is true "beyond a reasonable doubt." **Nor must the applicant provide evidence such that it establishes an asserted utility as a matter of statistical certainty.** Instead, evidence will be sufficient if, considered as a whole, it leads a person of ordinary skill in the art to conclude that the asserted utility is more likely than not. M.P.E.P. at § 2107.02, part VII (2004) (underline emphasis in original, bold emphasis added, internal citations omitted).

The PTO has the initial burden to offer evidence "that one of ordinary skill in the art would reasonably doubt the asserted utility." *In re Brana*, 51 F.3d 1560, 1566, 34 U.S.P.Q.2d 1436 (Fed. Cir. 1995). Only then does the burden shift to the Applicant to provide rebuttal

Appl. No. : 10/063,517  
Filed : May 7, 2002

evidence. *Id.* As stated in the M.P.E.P., such rebuttal evidence does not need to absolutely prove that the asserted utility is real. Rather, the evidence only needs to be reasonably indicative of the asserted utility.

Considering the totality of the evidence of record, Applicants submit that the PTO has not met its initial burden of overcoming the presumption that the asserted utility is sufficient to satisfy the utility requirement. The evidence proffered by the Examiner to counter Applicants' own evidence is either irrelevant to the issue at hand concerning the data in Example 18, (e.g., Pennica *et al.*, Hittleman *et al.*), or fails to establish that one skilled in the art would more likely than not doubt Applicants' asserted utility (e.g., Hu *et al.* and Gygi *et al.*). Further, even if the PTO has met its initial burden, the Applicants' supporting rebuttal evidence from experts and leading textbooks in the field is sufficient to establish that one of skill in the art would be more likely than not to believe that the claimed antibodies can be used as diagnostic tools for cancer, particularly lung cancer. Certainly, the Examiner has not provided evidence that establishes that Applicants' asserted utilities are "wholly inconsistent with the contemporary knowledge in the art." MPEP § 2107.02 III B.

### **Conclusion**

The Examiner has asserted several arguments for why there is a lack of a substantial utility: (1) that the data provided in Example 18 does not support the conclusion that the gene encoding PRO300 or PRO300 itself is differentially expressed in lung cancer compared to normal tissue; (2) that different types of lung tumor tissue would not be expected to have the same gene expression pattern; (3) that no mutation or translocation of PRO300 has been associated with lung cancer; (4) that aneuploid DNA can be found in normal tissue, and detection of increased DNA copy number does not necessarily correlate with a cancerous state; and (5) that Applicants have not provided evidentiary support for the proposition that increased mRNA levels are predictive of corresponding increased levels of the encoded polypeptide. For the above reasons, the Examiner has maintained the position that the claimed PRO300 antibodies cannot be used as cancer diagnostic or therapeutic tools. Applicants have addressed each of these arguments in turn.

First, the Applicants provided a first Declaration of Chris Grimaldi stating that given the at least two-fold difference in expression levels, the disclosed nucleic acids and corresponding polypeptides and antibodies have utility as cancer diagnostic tools. Applicants have also



Appl. No. : 10/063,517  
Filed : May 7, 2002

established that the asserted utilities for the claimed antibodies exist regardless of the cause or consequences of the differential expression. Next, Applicants have demonstrated that the possibility that PRO300 is not differentially expressed in all types of tumor does not detract from Applicants' asserted utility.

Finally, Applicants submit that the second Grimaldi Declaration and Polakis Declaration, the accompanying references, as well as the excerpts and references cited above, demonstrate that it is well-established in the art that a change in mRNA levels generally correlates to a corresponding change in the encoded protein levels. The Examiner has not offered any substantial reasoning or evidence to the contrary – they Examiner has cited to references that either do not relate to mRNA and protein expression levels, or that simply do not establish that one skilled in the art would reasonably doubt Applicants' asserted utility.

Given the totality of the evidence provided, Applicants submit that they have established a substantial, specific, and credible utility for the claimed antibodies as diagnostic tools. According to the PTO Utility Examination Guidelines (2001), irrefutable proof of a claimed utility is not required. Rather, a specific, substantial, and credible utility requires only a "reasonable" confirmation of a real world context of use. Applicants remind the PTO that:

A small degree of utility is sufficient . . . The claimed invention must only be capable of performing some beneficial function . . . An invention does not lack utility merely because the particular embodiment disclosed in the patent lacks perfection or performs crudely . . . A commercially successful product is not required . . . Nor is it essential that the invention accomplish all its intended functions . . . or operate under all conditions . . . partial success being sufficient to demonstrate patentable utility . . . In short, **the defense of non-utility cannot be sustained without proof of total incapacity**. If an invention is only partially successful in achieving a useful result, a rejection of the claimed invention as a whole based on a lack of utility is not appropriate. M.P.E.P. at 2107.01 (underline emphasis in original, bold emphasis added, citations omitted).

Applicants submit that they have established that it is more likely than not that one of skill in the art would reasonably accept the utility for the claimed antibodies as diagnostic tools as set forth in the specification. In view of the above, Applicants respectfully request that the Examiner reconsider and withdraw the utility rejection under 35 U.S.C. §101.

**Appl. No.** : 10/063,517  
**Filed** : May 7, 2002

**Rejection Under 35 U.S.C. § 112, First Paragraph – Enablement**

The Examiner has rejected Claims 1-5 under 35 U.S.C. §112, first paragraph, as not being enabled since the claimed invention allegedly lacks utility. According to the Examiner, significant further experimentation would be necessary to be able to use the claimed antibodies because no function has been determined for the same.

Applicants submit that in the discussion of the 35 U.S.C. § 101 rejection above, Applicants have established a substantial, specific, and credible utility for the claimed antibodies. Applicants therefore request that the Examiner reconsider and withdraw the enablement rejection to the extent that it is based on a lack of utility for the claimed antibodies.

**Rejection Under 35 U.S.C. § 102(b) – Anticipation**

The Examiner has maintained the rejection of Claims 1-8 and 11-13 as being anticipated by WO 01/16318 to Eaton et al., to which Applicants claim priority. According to the Examiner, the claimed invention does not fulfill the requirements of 35 U.S.C. § 112, and thus Applicants are not entitled to the benefit of WO 01/16318, and cites the same against Applicants' present application.

Under 35 U.S.C. § 120, an applicant is entitled to the benefit of the filing date of an earlier filed application that discloses the same invention in the manner provided by 35 U.S.C. § 112, first paragraph, provided the applicant properly claims priority to the earlier application. In a preliminary amendment filed on September 3, 2002, Applicants made specific reference to WO 01/16318, claiming priority thereto. WO 01/16318 contains the same disclosure relating to PRO300 and its utilities as the instant application, including the data in Example 18. For the same reasons detailed above in the Remarks addressed to the rejections under 35 U.S.C. § 101 and 112 in the instant response, Applicants submit that the WO 01/16318 is enabling for and adequately describes the claimed invention. Therefore, because Applicants have properly claimed priority to WO 01/16318, and because WO 01/16318 satisfies the requirements of 35 U.S.C. § 112, Applicants are fully entitled to the benefit of the filing date of WO 01/16318. Applicants respectfully request that the Examiner reconsider and withdraw the rejection under 35 U.S.C. § 102(b).

**Appl. No.** : 10/063,517  
**Filed** : May 7, 2002

### CONCLUSION

In view of the above, Applicants respectfully maintain that claims are patentable and request that they be passed to issue. Applicants invite the Examiner to call the undersigned if any remaining issues may be resolved by telephone.

Please charge any additional fees, including any fees for additional extension of time, or credit overpayment to Deposit Account No. 11-1410.

Respectfully submitted,

KNOBBE, MARTENS, OLSON & BEAR, LLP

Dated: May 26, 2005

By: Anne Marie Kaiser  
Anne Marie Kaiser  
Registration No. 37,649  
Attorney of Record  
Customer No. 30,313  
(619) 235-8550

1689541  
043005

**SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC ACIDS  
ENCODING THE SAME**

Related Applications

This application is a continuation of, and claims priority under 35 USC §120 to, US Application 10/006867 filed 12/6/2001, which is a continuation of, and claims priority under 35 USC §120 to, PCT Application PCT/US00/23328 filed 8/24/00, which is a continuation-in part of, and claims priority under 35 USC §120 to, US Application 09/380138 filed 8/25/1999, now abandoned, which is the National Stage filed under 35 USC §371 of PCT Application PCT/US99/05028 filed 3/8/1999, which claims priority under 35 USC §119 to US Provisional Application 60/082797 filed 4/22/1998.

Field of the Invention

**[0001]** The present invention relates generally to the identification and isolation of novel DNA and to the recombinant production of novel polypeptides.

Background of the Invention

**[0002]** Extracellular proteins play important roles in, among other things, the formation, differentiation and maintenance of multicellular organisms. The fate of many individual cells, e.g., proliferation, migration, differentiation, or interaction with other cells, is typically governed by information received from other cells and/or the immediate environment. This information is often transmitted by secreted polypeptides (for instance, mitogenic factors, survival factors, cytotoxic factors, differentiation factors, neuropeptides, and hormones) which are, in turn, received and interpreted by diverse cell receptors or membrane-bound proteins. These secreted polypeptides or signaling molecules normally pass through the cellular secretory pathway to reach their site of action in the extracellular environment.

**[0003]** Secreted proteins have various industrial applications, including as pharmaceuticals, diagnostics, biosensors and bioreactors. Most protein drugs available at present, such as thrombolytic agents, interferons, interleukins, erythropoietins, colony

stimulating factors, and various other cytokines, are secretory proteins. Their receptors, which are membrane proteins, also have potential as therapeutic or diagnostic agents. Efforts are being undertaken by both industry and academia to identify new, native secreted proteins. Many efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel secreted proteins. Examples of screening methods and techniques are described in the literature [see, for example, Klein et al., Proc. Natl. Acad. Sci. 93:7108-7113 (1996); U.S. Patent No. 5,536,637)].

**[0004]** Membrane-bound proteins and receptors can play important roles in, among other things, the formation, differentiation and maintenance of multicellular organisms. The fate of many individual cells, e.g., proliferation, migration, differentiation, or interaction with other cells, is typically governed by information received from other cells and/or the immediate environment. This information is often transmitted by secreted polypeptides (for instance, mitogenic factors, survival factors, cytotoxic factors, differentiation factors, neuropeptides, and hormones) which are, in turn, received and interpreted by diverse cell receptors or membrane-bound proteins. Such membrane-bound proteins and cell receptors include, but are not limited to, cytokine receptors, receptor kinases, receptor phosphatases, receptors involved in cell-cell interactions, and cellular adhesion molecules like selectins and integrins. For instance, transduction of signals that regulate cell growth and differentiation is regulated in part by phosphorylation of various cellular proteins. Protein tyrosine kinases, enzymes that catalyze that process, can also act as growth factor receptors. Examples include fibroblast growth factor receptor and nerve growth factor receptor.

**[0005]** Membrane-bound proteins and receptor molecules have various industrial applications, including as pharmaceutical and diagnostic agents. Receptor immunoadhesins, for instance, can be employed as therapeutic agents to block receptor-ligand interactions. The membrane-bound proteins can also be employed for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interaction.

**[0006]** Efforts are being undertaken by both industry and academia to identify new, native receptor or membrane-bound proteins. Many efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel receptor or membrane-bound proteins.

### Summary of the Invention

[0007] In one embodiment, the invention provides an isolated nucleic acid molecule comprising a nucleotide sequence that encodes a PRO polypeptide.

[0008] In one aspect, the isolated nucleic acid molecule comprises a nucleotide sequence having at least about 80% nucleic acid sequence identity, alternatively at least about 81% nucleic acid sequence identity, alternatively at least about 82% nucleic acid sequence identity, alternatively at least about 83% nucleic acid sequence identity, alternatively at least about 84% nucleic acid sequence identity, alternatively at least about 85% nucleic acid sequence identity, alternatively at least about 86% nucleic acid sequence identity, alternatively at least about 87% nucleic acid sequence identity, alternatively at least about 88% nucleic acid sequence identity, alternatively at least about 89% nucleic acid sequence identity, alternatively at least about 90% nucleic acid sequence identity, alternatively at least about 91% nucleic acid sequence identity, alternatively at least about 92% nucleic acid sequence identity, alternatively at least about 93% nucleic acid sequence identity, alternatively at least about 94% nucleic acid sequence identity, alternatively at least about 95% nucleic acid sequence identity, alternatively at least about 96% nucleic acid sequence identity, alternatively at least about 97% nucleic acid sequence identity, alternatively at least about 98% nucleic acid sequence identity and alternatively at least about 99% nucleic acid sequence identity to (a) a DNA molecule encoding a PRO polypeptide having a full-length amino acid sequence as disclosed herein, an amino acid sequence lacking the signal peptide as disclosed herein, an extracellular domain of a transmembrane protein, with or without the signal peptide, as disclosed herein or any other specifically defined fragment of the full-length amino acid sequence as disclosed herein, or (b) the complement of the DNA molecule of (a).

[0009] In other aspects, the isolated nucleic acid molecule comprises a nucleotide sequence having at least about 80% nucleic acid sequence identity, alternatively at least about 81% nucleic acid sequence identity, alternatively at least about 82% nucleic acid sequence identity, alternatively at least about 83% nucleic acid sequence identity, alternatively at least about 84% nucleic acid sequence identity, alternatively at least about 85% nucleic acid sequence identity, alternatively at least about 86% nucleic acid sequence identity, alternatively

at least about 87% nucleic acid sequence identity, alternatively at least about 88% nucleic acid sequence identity, alternatively at least about 89% nucleic acid sequence identity, alternatively at least about 90% nucleic acid sequence identity, alternatively at least about 91% nucleic acid sequence identity, alternatively at least about 92% nucleic acid sequence identity, alternatively at least about 93% nucleic acid sequence identity, alternatively at least about 94% nucleic acid sequence identity, alternatively at least about 95% nucleic acid sequence identity, alternatively at least about 96% nucleic acid sequence identity, alternatively at least about 97% nucleic acid sequence identity, alternatively at least about 98% nucleic acid sequence identity and alternatively at least about 99% nucleic acid sequence identity to (a) a DNA molecule comprising the coding sequence of a full-length PRO polypeptide cDNA as disclosed herein, the coding sequence of a PRO polypeptide lacking the signal peptide as disclosed herein, the coding sequence of an extracellular domain of a transmembrane PRO polypeptide, with or without the signal peptide, as disclosed herein or the coding sequence of any other specifically defined fragment of the full-length amino acid sequence as disclosed herein, or (b) the complement of the DNA molecule of (a).

[0010] In a further aspect, the invention concerns an isolated nucleic acid molecule comprising a nucleotide sequence having at least about 80% nucleic acid sequence identity, alternatively at least about 81% nucleic acid sequence identity, alternatively at least about 82% nucleic acid sequence identity, alternatively at least about 83% nucleic acid sequence identity, alternatively at least about 84% nucleic acid sequence identity, alternatively at least about 85% nucleic acid sequence identity, alternatively at least about 86% nucleic acid sequence identity, alternatively at least about 87% nucleic acid sequence identity, alternatively at least about 88% nucleic acid sequence identity, alternatively at least about 89% nucleic acid sequence identity, alternatively at least about 90% nucleic acid sequence identity, alternatively at least about 91% nucleic acid sequence identity, alternatively at least about 92% nucleic acid sequence identity, alternatively at least about 93% nucleic acid sequence identity, alternatively at least about 94% nucleic acid sequence identity, alternatively at least about 95% nucleic acid sequence identity, alternatively at least about 96% nucleic acid sequence identity, alternatively at least about 97% nucleic acid sequence identity, alternatively at least about 98% nucleic acid sequence identity and alternatively at least about 99% nucleic acid sequence identity to (a) a DNA molecule that

encodes the same mature polypeptide encoded by any of the human protein cDNAs deposited with the ATCC as disclosed herein, or (b) the complement of the DNA molecule of (a).

[0011] Another aspect the invention provides an isolated nucleic acid molecule comprising a nucleotide sequence encoding a PRO polypeptide which is either transmembrane domain-deleted or transmembrane domain-inactivated, or is complementary to such encoding nucleotide sequence, wherein the transmembrane domain(s) of such polypeptide are disclosed herein. Therefore, soluble extracellular domains of the herein described PRO polypeptides are contemplated.

[0012] Another embodiment is directed to fragments of a PRO polypeptide coding sequence, or the complement thereof, that may find use as, for example, hybridization probes, for encoding fragments of a PRO polypeptide that may optionally encode a polypeptide comprising a binding site for an anti-PRO antibody or as antisense oligonucleotide probes. Such nucleic acid fragments are usually at least about 20 nucleotides in length, alternatively at least about 30 nucleotides in length, alternatively at least about 40 nucleotides in length, alternatively at least about 50 nucleotides in length, alternatively at least about 60 nucleotides in length, alternatively at least about 70 nucleotides in length, alternatively at least about 80 nucleotides in length, alternatively at least about 90 nucleotides in length, alternatively at least about 100 nucleotides in length, alternatively at least about 110 nucleotides in length, alternatively at least about 120 nucleotides in length, alternatively at least about 130 nucleotides in length, alternatively at least about 140 nucleotides in length, alternatively at least about 150 nucleotides in length, alternatively at least about 160 nucleotides in length, alternatively at least about 170 nucleotides in length, alternatively at least about 180 nucleotides in length, alternatively at least about 190 nucleotides in length, alternatively at least about 200 nucleotides in length, alternatively at least about 250 nucleotides in length, alternatively at least about 300 nucleotides in length, alternatively at least about 350 nucleotides in length, alternatively at least about 400 nucleotides in length, alternatively at least about 450 nucleotides in length, alternatively at least about 500 nucleotides in length, alternatively at least about 600 nucleotides in length, alternatively at least about 700 nucleotides in length, alternatively at least about 800 nucleotides in length, alternatively at least about 900 nucleotides in length and alternatively at least about 1000 nucleotides in length,



wherein in this context the term “about” means the referenced nucleotide sequence length plus or minus 10% of that referenced length. It is noted that novel fragments of a PRO polypeptide-encoding nucleotide sequence may be determined in a routine manner by aligning the PRO polypeptide-encoding nucleotide sequence with other known nucleotide sequences using any of a number of well known sequence alignment programs and determining which PRO polypeptide-encoding nucleotide sequence fragment(s) are novel. All of such PRO polypeptide-encoding nucleotide sequences are contemplated herein. Also contemplated are the PRO polypeptide fragments encoded by these nucleotide molecule fragments, preferably those PRO polypeptide fragments that comprise a binding site for an anti-PRO antibody.

[0013] In another embodiment, the invention provides isolated PRO polypeptide encoded by any of the isolated nucleic acid sequences hereinabove identified.

[0014] In a certain aspect, the invention concerns an isolated PRO polypeptide, comprising an amino acid sequence having at least about 80% amino acid sequence identity, alternatively at least about 81% amino acid sequence identity, alternatively at least about 82% amino acid sequence identity, alternatively at least about 83% amino acid sequence identity, alternatively at least about 84% amino acid sequence identity, alternatively at least about 85% amino acid sequence identity, alternatively at least about 86% amino acid sequence identity, alternatively at least about 87% amino acid sequence identity, alternatively at least about 88% amino acid sequence identity, alternatively at least about 89% amino acid sequence identity, alternatively at least about 90% amino acid sequence identity, alternatively at least about 91% amino acid sequence identity, alternatively at least about 92% amino acid sequence identity, alternatively at least about 93% amino acid sequence identity, alternatively at least about 94% amino acid sequence identity, alternatively at least about 95% amino acid sequence identity, alternatively at least about 96% amino acid sequence identity, alternatively at least about 97% amino acid sequence identity, alternatively at least about 98% amino acid sequence identity and alternatively at least about 99% amino acid sequence identity to a PRO polypeptide having a full-length amino acid sequence as disclosed herein, an amino acid sequence lacking the signal peptide as disclosed herein, an extracellular domain of a transmembrane protein, with or without the signal peptide, as disclosed herein or any other specifically defined fragment of the full-length amino acid sequence as disclosed herein.

**[0015]** In a further aspect, the invention concerns an isolated PRO polypeptide comprising an amino acid sequence having at least about 80% amino acid sequence identity, alternatively at least about 81% amino acid sequence identity, alternatively at least about 82% amino acid sequence identity, alternatively at least about 83% amino acid sequence identity, alternatively at least about 84% amino acid sequence identity, alternatively at least about 85% amino acid sequence identity, alternatively at least about 86% amino acid sequence identity, alternatively at least about 87% amino acid sequence identity, alternatively at least about 88% amino acid sequence identity, alternatively at least about 89% amino acid sequence identity, alternatively at least about 90% amino acid sequence identity, alternatively at least about 91% amino acid sequence identity, alternatively at least about 92% amino acid sequence identity, alternatively at least about 93% amino acid sequence identity, alternatively at least about 94% amino acid sequence identity, alternatively at least about 95% amino acid sequence identity, alternatively at least about 96% amino acid sequence identity, alternatively at least about 97% amino acid sequence identity, alternatively at least about 98% amino acid sequence identity and alternatively at least about 99% amino acid sequence identity to an amino acid sequence encoded by any of the human protein cDNAs deposited with the ATCC as disclosed herein.

**[0016]** In a specific aspect, the invention provides an isolated PRO polypeptide without the N-terminal signal sequence and/or the initiating methionine and is encoded by a nucleotide sequence that encodes such an amino acid sequence as hereinbefore described. Processes for producing the same are also herein described, wherein those processes comprise culturing a host cell comprising a vector which comprises the appropriate encoding nucleic acid molecule under conditions suitable for expression of the PRO polypeptide and recovering the PRO polypeptide from the cell culture.

**[0017]** Another aspect the invention provides an isolated PRO polypeptide which is either transmembrane domain-deleted or transmembrane domain-inactivated. Processes for producing the same are also herein described, wherein those processes comprise culturing a host cell comprising a vector which comprises the appropriate encoding nucleic acid molecule under conditions suitable for expression of the PRO polypeptide and recovering the PRO polypeptide from the cell culture.

**[0018]** In yet another embodiment, the invention concerns agonists and antagonists of a native PRO polypeptide as defined herein. In a particular embodiment, the agonist or antagonist is an anti-PRO antibody or a small molecule.

**[0019]** In a further embodiment, the invention concerns a method of identifying agonists or antagonists to a PRO polypeptide which comprise contacting the PRO polypeptide with a candidate molecule and monitoring a biological activity mediated by said PRO polypeptide. Preferably, the PRO polypeptide is a native PRO polypeptide.

**[0020]** In a still further embodiment, the invention concerns a composition of matter comprising a PRO polypeptide, or an agonist or antagonist of a PRO polypeptide as herein described, or an anti-PRO antibody, in combination with a carrier. Optionally, the carrier is a pharmaceutically acceptable carrier.

**[0021]** Another embodiment of the present invention is directed to the use of a PRO polypeptide, or an agonist or antagonist thereof as hereinbefore described, or an anti-PRO antibody, for the preparation of a medicament useful in the treatment of a condition which is responsive to the PRO polypeptide, an agonist or antagonist thereof or an anti-PRO antibody.

**[0022]** In other embodiments of the present invention, the invention provides vectors comprising DNA encoding any of the herein described polypeptides. Host cell comprising any such vector are also provided. By way of example, the host cells may be CHO cells, *E. coli*, or yeast. A process for producing any of the herein described polypeptides is further provided and comprises culturing host cells under conditions suitable for expression of the desired polypeptide and recovering the desired polypeptide from the cell culture.

**[0023]** In other embodiments, the invention provides chimeric molecules comprising any of the herein described polypeptides fused to a heterologous polypeptide or amino acid sequence. Example of such chimeric molecules comprise any of the herein described polypeptides fused to an epitope tag sequence or a Fc region of an immunoglobulin.

**[0024]** In another embodiment, the invention provides an antibody which binds, preferably specifically, to any of the above or below described polypeptides. Optionally, the antibody is a monoclonal antibody, humanized antibody, antibody fragment or single-chain antibody.

[0025] In yet other embodiments, the invention provides oligonucleotide probes useful for isolating genomic and cDNA nucleotide sequences or as antisense probes, wherein those probes may be derived from any of the above or below described nucleotide sequences.

[0026] In yet other embodiments, the present invention is directed to methods of using the PRO polypeptides of the present invention for a variety of uses based upon the functional biological assay data presented in the Examples below.

#### Brief Description of the Drawings

[0027] Figure 1 shows a nucleotide sequence (SEQ ID NO:1) of a native sequence PRO180 cDNA, wherein SEQ ID NO:1 is a clone designated herein as “DNA26843-1389”.

[0028] Figure 2 shows the amino acid sequence (SEQ ID NO:2) derived from the coding sequence of SEQ ID NO:1 shown in Figure 1.

[0029] Figure 3 shows a nucleotide sequence (SEQ ID NO:3) of a native sequence PRO218 cDNA, wherein SEQ ID NO:3 is a clone designated herein as “DNA30867-1335”.

[0030] Figure 4 shows the amino acid sequence (SEQ ID NO:4) derived from the coding sequence of SEQ ID NO:3 shown in Figure 3.

[0031] Figure 5 shows a nucleotide sequence (SEQ ID NO:5) of a native sequence PRO263 cDNA, wherein SEQ ID NO:5 is a clone designated herein as “DNA34431-1177”.

[0032] Figure 6 shows the amino acid sequence (SEQ ID NO:6) derived from the coding sequence of SEQ ID NO:5 shown in Figure 5.

[0033] Figure 7 shows a nucleotide sequence (SEQ ID NO:7) of a native sequence PRO295 cDNA, wherein SEQ ID NO:7 is a clone designated herein as “DNA38268-1188”.

[0034] Figure 8 shows the amino acid sequence (SEQ ID NO:8) derived from the coding sequence of SEQ ID NO:7 shown in Figure 7.

[0035] Figure 9 shows a nucleotide sequence (SEQ ID NO:9) of a native sequence PRO874 cDNA, wherein SEQ ID NO:9 is a clone designated herein as “DNA40621-1440”.

[0036] Figure 10 shows the amino acid sequence (SEQ ID NO:10) derived from the coding sequence of SEQ ID NO:9 shown in Figure 9.

[0037] Figure 11 shows a nucleotide sequence (SEQ ID NO:11) of a native sequence PRO300 cDNA, wherein SEQ ID NO:11 is a clone designated herein as “DNA40625-1189”.

[0038] Figure 12 shows the amino acid sequence (SEQ ID NO:12) derived from the coding sequence of SEQ ID NO:11 shown in Figure 11.

[0039] Figure 13 shows a nucleotide sequence (SEQ ID NO:13) of a native sequence PRO1864 cDNA, wherein SEQ ID NO:13 is a clone designated herein as “DNA45409-2511”.

[0040] Figure 14 shows the amino acid sequence (SEQ ID NO:14) derived from the coding sequence of SEQ ID NO:13 shown in Figure 13.

[0041] Figure 15 shows a nucleotide sequence (SEQ ID NO:15) of a native sequence PRO1282 cDNA, wherein SEQ ID NO:15 is a clone designated herein as “DNA45495-1550”.

[0042] Figure 16 shows the amino acid sequence (SEQ ID NO:16) derived from the coding sequence of SEQ ID NO:15 shown in Figure 15.

[0043] Figure 17 shows a nucleotide sequence (SEQ ID NO:17) of a native sequence PRO1063 cDNA, wherein SEQ ID NO:17 is a clone designated herein as “DNA49820-1427”.

[0044] Figure 18 shows the amino acid sequence (SEQ ID NO:18) derived from the coding sequence of SEQ ID NO:17 shown in Figure 17.

[0045] Figure 19 shows a nucleotide sequence (SEQ ID NO:19) of a native sequence PRO1773 cDNA, wherein SEQ ID NO:19 is a clone designated herein as “DNA56406-1704”.

[0046] Figure 20 shows the amino acid sequence (SEQ ID NO:20) derived from the coding sequence of SEQ ID NO:19 shown in Figure 19.

[0047] Figure 21 shows a nucleotide sequence (SEQ ID NO:21) of a native sequence PRO1013 cDNA, wherein SEQ ID NO:21 is a clone designated herein as “DNA56410-1414”.

[0048] Figure 22 shows the amino acid sequence (SEQ ID NO:22) derived from the coding sequence of SEQ ID NO:21 shown in Figure 21.

**[0049]** Figure 23 shows a nucleotide sequence (SEQ ID NO:23) of a native sequence PRO937 cDNA, wherein SEQ ID NO:23 is a clone designated herein as “DNA56436-1448”.

**[0050]** Figure 24 shows the amino acid sequence (SEQ ID NO:24) derived from the coding sequence of SEQ ID NO:23 shown in Figure 23.

**[0051]** Figure 25 shows a nucleotide sequence (SEQ ID NO:25) of a native sequence PRO842 cDNA, wherein SEQ ID NO:25 is a clone designated herein as “DNA56855-1447”.

**[0052]** Figure 26 shows the amino acid sequence (SEQ ID NO:26) derived from the coding sequence of SEQ ID NO:25 shown in Figure 25.

**[0053]** Figure 27 shows a nucleotide sequence (SEQ ID NO:27) of a native sequence PRO1180 cDNA, wherein SEQ ID NO:27 is a clone designated herein as “DNA56860-1510”.

**[0054]** Figure 28 shows the amino acid sequence (SEQ ID NO:28) derived from the coding sequence of SEQ ID NO:27 shown in Figure 27.

**[0055]** Figure 29 shows a nucleotide sequence (SEQ ID NO:29) of a native sequence PRO831 cDNA, wherein SEQ ID NO:29 is a clone designated herein as “DNA56862-1343”.

**[0056]** Figure 30 shows the amino acid sequence (SEQ ID NO:30) derived from the coding sequence of SEQ ID NO:29 shown in Figure 29.

**[0057]** Figure 31 shows a nucleotide sequence (SEQ ID NO:31) of a native sequence PRO1115 cDNA, wherein SEQ ID NO:31 is a clone designated herein as “DNA56868-1478”.

**[0058]** Figure 32 shows the amino acid sequence (SEQ ID NO:32) derived from the coding sequence of SEQ ID NO:31 shown in Figure 31.

**[0059]** Figure 33 shows a nucleotide sequence (SEQ ID NO:33) of a native sequence PRO1277 cDNA, wherein SEQ ID NO:33 is a clone designated herein as “DNA56869-1545”.

**[0060]** Figure 34 shows the amino acid sequence (SEQ ID NO:34) derived from the coding sequence of SEQ ID NO:33 shown in Figure 33.

**[0061]** Figure 35 shows a nucleotide sequence (SEQ ID NO:35) of a native sequence PRO1074 cDNA, wherein SEQ ID NO:35 is a clone designated herein as “DNA57704-1452”.

**[0062]** Figure 36 shows the amino acid sequence (SEQ ID NO:36) derived from the coding sequence of SEQ ID NO:35 shown in Figure 35.

**[0063]** Figure 37 shows a nucleotide sequence (SEQ ID NO:37) of a native sequence PRO1344 cDNA, wherein SEQ ID NO:37 is a clone designated herein as “DNA58723-1588”.

**[0064]** Figure 38 shows the amino acid sequence (SEQ ID NO:38) derived from the coding sequence of SEQ ID NO:37 shown in Figure 37.

**[0065]** Figure 39 shows a nucleotide sequence (SEQ ID NO:39) of a native sequence PRO1136 cDNA, wherein SEQ ID NO:39 is a clone designated herein as “DNA57827-1493”.

**[0066]** Figure 40 shows the amino acid sequence (SEQ ID NO:40) derived from the coding sequence of SEQ ID NO:39 shown in Figure 39.

**[0067]** Figure 41 shows a nucleotide sequence (SEQ ID NO:41) of a native sequence PRO1109 cDNA, wherein SEQ ID NO:41 is a clone designated herein as “DNA58737-1473”.

**[0068]** Figure 42 shows the amino acid sequence (SEQ ID NO:42) derived from the coding sequence of SEQ ID NO:41 shown in Figure 41.

**[0069]** Figure 43 shows a nucleotide sequence (SEQ ID NO:43) of a native sequence PRO1003 cDNA, wherein SEQ ID NO:43 is a clone designated herein as “DNA58846-1409”.

**[0070]** Figure 44 shows the amino acid sequence (SEQ ID NO:44) derived from the coding sequence of SEQ ID NO:43 shown in Figure 43.

**[0071]** Figure 45 shows a nucleotide sequence (SEQ ID NO:45) of a native sequence PRO1138 cDNA, wherein SEQ ID NO:45 is a clone designated herein as “DNA58850-1495”.

**[0072]** Figure 46 shows the amino acid sequence (SEQ ID NO:46) derived from the coding sequence of SEQ ID NO:45 shown in Figure 45.

**[0073]** Figure 47 shows a nucleotide sequence (SEQ ID NO:47) of a native sequence PRO994 cDNA, wherein SEQ ID NO:47 is a clone designated herein as “DNA58855-1422”.

**[0074]** Figure 48 shows the amino acid sequence (SEQ ID NO:48) derived from the coding sequence of SEQ ID NO:47 shown in Figure 47.

**[0075]** Figure 49 shows a nucleotide sequence (SEQ ID NO:49) of a native sequence PRO1069 cDNA, wherein SEQ ID NO:49 is a clone designated herein as “DNA59211-1450”.

**[0076]** Figure 50 shows the amino acid sequence (SEQ ID NO:50) derived from the coding sequence of SEQ ID NO:49 shown in Figure 49.

**[0077]** Figure 51 shows a nucleotide sequence (SEQ ID NO:51) of a native sequence PRO1411 cDNA, wherein SEQ ID NO:51 is a clone designated herein as “DNA59212-1627”.

**[0078]** Figure 52 shows the amino acid sequence (SEQ ID NO:52) derived from the coding sequence of SEQ ID NO:51 shown in Figure 51.

**[0079]** Figure 53 shows a nucleotide sequence (SEQ ID NO:53) of a native sequence PRO1129 cDNA, wherein SEQ ID NO:53 is a clone designated herein as “DNA59213-1487”.

**[0080]** Figure 54 shows the amino acid sequence (SEQ ID NO:54) derived from the coding sequence of SEQ ID NO:53 shown in Figure 53.

**[0081]** Figure 55 shows a nucleotide sequence (SEQ ID NO:55) of a native sequence PRO1027 cDNA, wherein SEQ ID NO:55 is a clone designated herein as “DNA59605-1418”.

**[0082]** Figure 56 shows the amino acid sequence (SEQ ID NO:56) derived from the coding sequence of SEQ ID NO:55 shown in Figure 55.

**[0083]** Figure 57 shows a nucleotide sequence (SEQ ID NO:57) of a native sequence PRO1106 cDNA, wherein SEQ ID NO:57 is a clone designated herein as “DNA59609-1470”.

**[0084]** Figure 58 shows the amino acid sequence (SEQ ID NO:58) derived from the coding sequence of SEQ ID NO:57 shown in Figure 57.



[0085] Figure 59 shows a nucleotide sequence (SEQ ID NO:59) of a native sequence PRO1291 cDNA, wherein SEQ ID NO:59 is a clone designated herein as “DNA59610-1556”.

[0086] Figure 60 shows the amino acid sequence (SEQ ID NO:60) derived from the coding sequence of SEQ ID NO:59 shown in Figure 59.

[0087] Figure 61 shows a nucleotide sequence (SEQ ID NO:61) of a native sequence PRO3573 cDNA, wherein SEQ ID NO:61 is a clone designated herein as “DNA59837-2545”.

[0088] Figure 62 shows the amino acid sequence (SEQ ID NO:62) derived from the coding sequence of SEQ ID NO:61 shown in Figure 61.

[0089] Figure 63 shows a nucleotide sequence (SEQ ID NO:63) of a native sequence PRO3566 cDNA, wherein SEQ ID NO:63 is a clone designated herein as “DNA59844-2542”.

[0090] Figure 64 shows the amino acid sequence (SEQ ID NO:64) derived from the coding sequence of SEQ ID NO:63 shown in Figure 63.

[0091] Figure 65 shows a nucleotide sequence (SEQ ID NO:65) of a native sequence PRO1098 cDNA, wherein SEQ ID NO:65 is a clone designated herein as “DNA59854-1459”.

[0092] Figure 66 shows the amino acid sequence (SEQ ID NO:66) derived from the coding sequence of SEQ ID NO:65 shown in Figure 65.

[0093] Figure 67 shows a nucleotide sequence (SEQ ID NO:67) of a native sequence PRO1158 cDNA, wherein SEQ ID NO:67 is a clone designated herein as “DNA60625-1507”.

[0094] Figure 68 shows the amino acid sequence (SEQ ID NO:68) derived from the coding sequence of SEQ ID NO:67 shown in Figure 67.

[0095] Figure 69 shows a nucleotide sequence (SEQ ID NO:69) of a native sequence PRO1124 cDNA, wherein SEQ ID NO:69 is a clone designated herein as “DNA60629-1481”.

[0096] Figure 70 shows the amino acid sequence (SEQ ID NO:70) derived from the coding sequence of SEQ ID NO:69 shown in Figure 69.

**[0097]** Figure 71 shows a nucleotide sequence (SEQ ID NO:71) of a native sequence PRO1287 cDNA, wherein SEQ ID NO:71 is a clone designated herein as “DNA61755-1554”.

**[0098]** Figure 72 shows the amino acid sequence (SEQ ID NO:72) derived from the coding sequence of SEQ ID NO:71 shown in Figure 71.

**[0099]** Figure 73 shows a nucleotide sequence (SEQ ID NO:73) of a native sequence PRO1335 cDNA, wherein SEQ ID NO:73 is a clone designated herein as “DNA62812-1594”.

**[0100]** Figure 74 shows the amino acid sequence (SEQ ID NO:74) derived from the coding sequence of SEQ ID NO:73 shown in Figure 73.

**[0101]** Figure 75 shows a nucleotide sequence (SEQ ID NO:75) of a native sequence PRO1315 cDNA, wherein SEQ ID NO:75 is a clone designated herein as “DNA62815-1576”.

**[0102]** Figure 76 shows the amino acid sequence (SEQ ID NO:76) derived from the coding sequence of SEQ ID NO:75 shown in Figure 75.

**[0103]** Figure 77 shows a nucleotide sequence (SEQ ID NO:77) of a native sequence PRO1357 cDNA, wherein SEQ ID NO:77 is a clone designated herein as “DNA64881-1602”.

**[0104]** Figure 78 shows the amino acid sequence (SEQ ID NO:78) derived from the coding sequence of SEQ ID NO:77 shown in Figure 77.

**[0105]** Figure 79 shows a nucleotide sequence (SEQ ID NO:79) of a native sequence PRO1356 cDNA, wherein SEQ ID NO:79 is a clone designated herein as “DNA64886-1601”.

**[0106]** Figure 80 shows the amino acid sequence (SEQ ID NO:80) derived from the coding sequence of SEQ ID NO:79 shown in Figure 79.

**[0107]** Figure 81 shows a nucleotide sequence (SEQ ID NO:81) of a native sequence PRO1557 cDNA, wherein SEQ ID NO:81 is a clone designated herein as “DNA64902-1667”.

**[0108]** Figure 82 shows the amino acid sequence (SEQ ID NO:82) derived from the coding sequence of SEQ ID NO:81 shown in Figure 81.

**[0109]** Figure 83 shows a nucleotide sequence (SEQ ID NO:83) of a native sequence PRO1347 cDNA, wherein SEQ ID NO:83 is a clone designated herein as “DNA64950-1590”.

**[0110]** Figure 84 shows the amino acid sequence (SEQ ID NO:84) derived from the coding sequence of SEQ ID NO:83 shown in Figure 83.

**[0111]** Figure 85 shows a nucleotide sequence (SEQ ID NO:85) of a native sequence PRO1302 cDNA, wherein SEQ ID NO:85 is a clone designated herein as “DNA65403-1565”.

**[0112]** Figure 86 shows the amino acid sequence (SEQ ID NO:86) derived from the coding sequence of SEQ ID NO:85 shown in Figure 85.

**[0113]** Figure 87 shows a nucleotide sequence (SEQ ID NO:87) of a native sequence PRO1270 cDNA, wherein SEQ ID NO:87 is a clone designated herein as “DNA66308-1537”.

**[0114]** Figure 88 shows the amino acid sequence (SEQ ID NO:88) derived from the coding sequence of SEQ ID NO:87 shown in Figure 87.

**[0115]** Figure 89 shows a nucleotide sequence (SEQ ID NO:89) of a native sequence PRO1268 cDNA, wherein SEQ ID NO:89 is a clone designated herein as “DNA66519-1535”.

**[0116]** Figure 90 shows the amino acid sequence (SEQ ID NO:90) derived from the coding sequence of SEQ ID NO:89 shown in Figure 89.

**[0117]** Figure 91 shows a nucleotide sequence (SEQ ID NO:91) of a native sequence PRO1327 cDNA, wherein SEQ ID NO:91 is a clone designated herein as “DNA66521-1583”.

**[0118]** Figure 92 shows the amino acid sequence (SEQ ID NO:92) derived from the coding sequence of SEQ ID NO:91 shown in Figure 91.

**[0119]** Figure 93 shows a nucleotide sequence (SEQ ID NO:93) of a native sequence PRO1328 cDNA, wherein SEQ ID NO:93 is a clone designated herein as “DNA66658-1584”.

**[0120]** Figure 94 shows the amino acid sequence (SEQ ID NO:94) derived from the coding sequence of SEQ ID NO:93 shown in Figure 93.

**[0121]** Figure 95 shows a nucleotide sequence (SEQ ID NO:95) of a native sequence PRO1329 cDNA, wherein SEQ ID NO:95 is a clone designated herein as “DNA66660-1585”.

**[0122]** Figure 96 shows the amino acid sequence (SEQ ID NO:96) derived from the coding sequence of SEQ ID NO:95 shown in Figure 95.

**[0123]** Figure 97 shows a nucleotide sequence (SEQ ID NO:97) of a native sequence PRO1340 cDNA, wherein SEQ ID NO:97 is a clone designated herein as “DNA66663-1598”.

**[0124]** Figure 98 shows the amino acid sequence (SEQ ID NO:98) derived from the coding sequence of SEQ ID NO:97 shown in Figure 97.

**[0125]** Figure 99 shows a nucleotide sequence (SEQ ID NO:99) of a native sequence PRO1342 cDNA, wherein SEQ ID NO:99 is a clone designated herein as “DNA66674-1599”.

**[0126]** Figure 100 shows the amino acid sequence (SEQ ID NO:100) derived from the coding sequence of SEQ ID NO:99 shown in Figure 99.

**[0127]** Figure 101 shows a nucleotide sequence (SEQ ID NO:101) of a native sequence PRO3579 cDNA, wherein SEQ ID NO:101 is a clone designated herein as “DNA68862-2546”.

**[0128]** Figure 102 shows the amino acid sequence (SEQ ID NO:102) derived from the coding sequence of SEQ ID NO:101 shown in Figure 101.

**[0129]** Figure 103 shows a nucleotide sequence (SEQ ID NO:103) of a native sequence PRO1472 cDNA, wherein SEQ ID NO:103 is a clone designated herein as “DNA68866-1644”.

**[0130]** Figure 104 shows the amino acid sequence (SEQ ID NO:104) derived from the coding sequence of SEQ ID NO:103 shown in Figure 103.

**[0131]** Figure 105 shows a nucleotide sequence (SEQ ID NO:105) of a native sequence PRO1461 cDNA, wherein SEQ ID NO:105 is a clone designated herein as “DNA68871-1638”.

**[0132]** Figure 106 shows the amino acid sequence (SEQ ID NO:106) derived from the coding sequence of SEQ ID NO:105 shown in Figure 105.

**[0133]** Figure 107 shows a nucleotide sequence (SEQ ID NO:107) of a native sequence PRO1568 cDNA, wherein SEQ ID NO:107 is a clone designated herein as “DNA68880-1676”.

**[0134]** Figure 108 shows the amino acid sequence (SEQ ID NO:108) derived from the coding sequence of SEQ ID NO:107 shown in Figure 107.

**[0135]** Figure 109 shows a nucleotide sequence (SEQ ID NO:109) of a native sequence PRO1753 cDNA, wherein SEQ ID NO:109 is a clone designated herein as “DNA68883-1691”.

**[0136]** Figure 110 shows the amino acid sequence (SEQ ID NO:110) derived from the coding sequence of SEQ ID NO:109 shown in Figure 109.

**[0137]** Figure 111 shows a nucleotide sequence (SEQ ID NO:111) of a native sequence PRO1570 cDNA, wherein SEQ ID NO:111 is a clone designated herein as “DNA68885-1678”.

**[0138]** Figure 112 shows the amino acid sequence (SEQ ID NO:112) derived from the coding sequence of SEQ ID NO:111 shown in Figure 111.

**[0139]** Figure 113 shows a nucleotide sequence (SEQ ID NO:113) of a native sequence PRO1446 cDNA, wherein SEQ ID NO:113 is a clone designated herein as “DNA71277-1636”.

**[0140]** Figure 114 shows the amino acid sequence (SEQ ID NO:114) derived from the coding sequence of SEQ ID NO:113 shown in Figure 113.

**[0141]** Figure 115 shows a nucleotide sequence (SEQ ID NO:115) of a native sequence PRO1565 cDNA, wherein SEQ ID NO:115 is a clone designated herein as “DNA73727-1673”.

**[0142]** Figure 116 shows the amino acid sequence (SEQ ID NO:116) derived from the coding sequence of SEQ ID NO:115 shown in Figure 115.

**[0143]** Figure 117 shows a nucleotide sequence (SEQ ID NO:117) of a native sequence PRO1572 cDNA, wherein SEQ ID NO:117 is a clone designated herein as “DNA73734-1680”.

**[0144]** Figure 118 shows the amino acid sequence (SEQ ID NO:118) derived from the coding sequence of SEQ ID NO:117 shown in Figure 117.

**[0145]** Figure 119 shows a nucleotide sequence (SEQ ID NO:119) of a native sequence PRO1573 cDNA, wherein SEQ ID NO:119 is a clone designated herein as “DNA73735-1681”.

**[0146]** Figure 120 shows the amino acid sequence (SEQ ID NO:120) derived from the coding sequence of SEQ ID NO:119 shown in Figure 119.

**[0147]** Figure 121 shows a nucleotide sequence (SEQ ID NO:121) of a native sequence PRO1550 cDNA, wherein SEQ ID NO:121 is a clone designated herein as “DNA76393-1664”.

**[0148]** Figure 122 shows the amino acid sequence (SEQ ID NO:122) derived from the coding sequence of SEQ ID NO:121 shown in Figure 121.

**[0149]** Figure 123 shows a nucleotide sequence (SEQ ID NO:123) of a native sequence PRO1693 cDNA, wherein SEQ ID NO:123 is a clone designated herein as “DNA77301-1708”.

**[0150]** Figure 124 shows the amino acid sequence (SEQ ID NO:124) derived from the coding sequence of SEQ ID NO:123 shown in Figure 123.

**[0151]** Figure 125 shows a nucleotide sequence (SEQ ID NO:125) of a native sequence PRO1566 cDNA, wherein SEQ ID NO:125 is a clone designated herein as “DNA77568-1626”.

**[0152]** Figure 126 shows the amino acid sequence (SEQ ID NO:126) derived from the coding sequence of SEQ ID NO:125 shown in Figure 125.

**[0153]** Figure 127 shows a nucleotide sequence (SEQ ID NO:127) of a native sequence PRO1774 cDNA, wherein SEQ ID NO:127 is a clone designated herein as “DNA77626-1705”.

**[0154]** Figure 128 shows the amino acid sequence (SEQ ID NO:128) derived from the coding sequence of SEQ ID NO:127 shown in Figure 127.

**[0155]** Figure 129 shows a nucleotide sequence (SEQ ID NO:129) of a native sequence PRO1928 cDNA, wherein SEQ ID NO:129 is a clone designated herein as “DNA81754-2532”.

**[0156]** Figure 130 shows the amino acid sequence (SEQ ID NO:130) derived from the coding sequence of SEQ ID NO:129 shown in Figure 129.

**[0157]** Figure 131 shows a nucleotide sequence (SEQ ID NO:131) of a native sequence PRO1865 cDNA, wherein SEQ ID NO:131 is a clone designated herein as “DNA81757-2512”.

**[0158]** Figure 132 shows the amino acid sequence (SEQ ID NO:132) derived from the coding sequence of SEQ ID NO:131 shown in Figure 131.

**[0159]** Figure 133 shows a nucleotide sequence (SEQ ID NO:133) of a native sequence PRO1925 cDNA, wherein SEQ ID NO:133 is a clone designated herein as “DNA82302-2529”.

**[0160]** Figure 134 shows the amino acid sequence (SEQ ID NO:134) derived from the coding sequence of SEQ ID NO:133 shown in Figure 133.

**[0161]** Figure 135 shows a nucleotide sequence (SEQ ID NO:135) of a native sequence PRO1926 cDNA, wherein SEQ ID NO:135 is a clone designated herein as “DNA82340-2530”.

**[0162]** Figure 136 shows the amino acid sequence (SEQ ID NO:136) derived from the coding sequence of SEQ ID NO:135 shown in Figure 135.

**[0163]** Figure 137 shows a nucleotide sequence (SEQ ID NO:137) of a native sequence PRO1801 cDNA, wherein SEQ ID NO:137 is a clone designated herein as “DNA83500-2506”.

**[0164]** Figure 138 shows the amino acid sequence (SEQ ID NO:138) derived from the coding sequence of SEQ ID NO:137 shown in Figure 137.

**[0165]** Figure 139 shows a nucleotide sequence (SEQ ID NO:139) of a native sequence PRO4405 cDNA, wherein SEQ ID NO:139 is a clone designated herein as “DNA84920-2614”.

**[0166]** Figure 140 shows the amino acid sequence (SEQ ID NO:140) derived from the coding sequence of SEQ ID NO:139 shown in Figure 139.

**[0167]** Figure 141 shows a nucleotide sequence (SEQ ID NO:141) of a native sequence PRO3435 cDNA, wherein SEQ ID NO:141 is a clone designated herein as “DNA85066-2534”.

**[0168]** Figure 142 shows the amino acid sequence (SEQ ID NO:142) derived from the coding sequence of SEQ ID NO:141 shown in Figure 141.

**[0169]** Figure 143 shows a nucleotide sequence (SEQ ID NO:143) of a native sequence PRO3543 cDNA, wherein SEQ ID NO:143 is a clone designated herein as “DNA86571-2551”.

**[0170]** Figure 144 shows the amino acid sequence (SEQ ID NO:144) derived from the coding sequence of SEQ ID NO:143 shown in Figure 143.

**[0171]** Figure 145 shows a nucleotide sequence (SEQ ID NO:145) of a native sequence PRO3443 cDNA, wherein SEQ ID NO:145 is a clone designated herein as “DNA87991-2540”.

**[0172]** Figure 146 shows the amino acid sequence (SEQ ID NO:146) derived from the coding sequence of SEQ ID NO:145 shown in Figure 145.

**[0173]** Figure 147 shows a nucleotide sequence (SEQ ID NO:147) of a native sequence PRO3442 cDNA, wherein SEQ ID NO:147 is a clone designated herein as “DNA92238-2539”.

**[0174]** Figure 148 shows the amino acid sequence (SEQ ID NO:148) derived from the coding sequence of SEQ ID NO:147 shown in Figure 147.

**[0175]** Figure 149 shows a nucleotide sequence (SEQ ID NO:149) of a native sequence PRO5990 cDNA, wherein SEQ ID NO:149 is a clone designated herein as “DNA96042-2682”.

**[0176]** Figure 150 shows the amino acid sequence (SEQ ID NO:150) derived from the coding sequence of SEQ ID NO:149 shown in Figure 149.

**[0177]** Figure 151 shows a nucleotide sequence (SEQ ID NO:151) of a native sequence PRO4342 cDNA, wherein SEQ ID NO:151 is a clone designated herein as “DNA96787-2534”.

**[0178]** Figure 152 shows the amino acid sequence (SEQ ID NO:152) derived from the coding sequence of SEQ ID NO:151 shown in Figure 151.

**[0179]** Figure 153 shows a nucleotide sequence (SEQ ID NO:153) of a native sequence PRO10096 cDNA, wherein SEQ ID NO:153 is a clone designated herein as “DNA125185-2806”.

**[0180]** Figure 154 shows the amino acid sequence (SEQ ID NO:154) derived from the coding sequence of SEQ ID NO:153 shown in Figure 153.



**[0181]** Figure 155 shows a nucleotide sequence (SEQ ID NO:155) of a native sequence PRO10272 cDNA, wherein SEQ ID NO:155 is a clone designated herein as “DNA147531-2821”.

**[0182]** Figure 156 shows the amino acid sequence (SEQ ID NO:156) derived from the coding sequence of SEQ ID NO:155 shown in Figure 155.

**[0183]** Figure 157 shows a nucleotide sequence (SEQ ID NO:157) of a native sequence PRO5801 cDNA, wherein SEQ ID NO:157 is a clone designated herein as “DNA115291-2681”.

**[0184]** Figure 158 shows the amino acid sequence (SEQ ID NO:158) derived from the coding sequence of SEQ ID NO:157 shown in Figure 157.

**[0185]** Figure 159 shows a nucleotide sequence (SEQ ID NO:159) of a native sequence PRO20110 cDNA, wherein SEQ ID NO:159 is a clone designated herein as “DNA166819”.

**[0186]** Figure 160 shows the amino acid sequence (SEQ ID NO:160) derived from the coding sequence of SEQ ID NO:159 shown in Figure 159.

**[0187]** Figure 161 shows a nucleotide sequence (SEQ ID NO:161) of a native sequence PRO20040 cDNA, wherein SEQ ID NO:161 is a clone designated herein as “DNA164625-2890”.

**[0188]** Figure 162 shows the amino acid sequence (SEQ ID NO:162) derived from the coding sequence of SEQ ID NO:161 shown in Figure 161.

**[0189]** Figure 163 shows a nucleotide sequence (SEQ ID NO:163) of a native sequence PRO20233 cDNA, wherein SEQ ID NO:163 is a clone designated herein as “DNA165608”.

**[0190]** Figure 164 shows the amino acid sequence (SEQ ID NO:164) derived from the coding sequence of SEQ ID NO:163 shown in Figure 163.

**[0191]** Figure 165 shows a nucleotide sequence (SEQ ID NO:165) of a native sequence PRO19670 cDNA, wherein SEQ ID NO:165 is a clone designated herein as “DNA131639-2874”.

**[0192]** Figure 166 shows the amino acid sequence (SEQ ID NO:166) derived from the coding sequence of SEQ ID NO:165 shown in Figure 165.

[0193] Figure 167 shows a nucleotide sequence (SEQ ID NO:167) of a native sequence PRO1890 cDNA, wherein SEQ ID NO:167 is a clone designated herein as “DNA79230-2525”.

[0194] Figure 168 shows the amino acid sequence (SEQ ID NO:168) derived from the coding sequence of SEQ ID NO:167 shown in Figure 167.

### Detailed Description of the Preferred Embodiments

#### I. Definitions

[0195] The terms “PRO polypeptide” and “PRO” as used herein and when immediately followed by a numerical designation refer to various polypeptides, wherein the complete designation (i.e., PRO/number) refers to specific polypeptide sequences as described herein. The terms “PRO/number polypeptide” and “PRO/number” wherein the term “number” is provided as an actual numerical designation as used herein encompass native sequence polypeptides and polypeptide variants (which are further defined herein). The PRO polypeptides described herein may be isolated from a variety of sources, such as from human tissue types or from another source, or prepared by recombinant or synthetic methods. The term “PRO polypeptide” refers to each individual PRO/number polypeptide disclosed herein. All disclosures in this specification which refer to the “PRO polypeptide” refer to each of the polypeptides individually as well as jointly. For example, descriptions of the preparation of, purification of, derivation of, formation of antibodies to or against, administration of, compositions containing, treatment of a disease with, etc., pertain to each polypeptide of the invention individually. The term “PRO polypeptide” also includes variants of the PRO/number polypeptides disclosed herein.

[0196] A “native sequence PRO polypeptide” comprises a polypeptide having the same amino acid sequence as the corresponding PRO polypeptide derived from nature. Such native sequence PRO polypeptides can be isolated from nature or can be produced by recombinant or synthetic means. The term “native sequence PRO polypeptide” specifically encompasses naturally-occurring truncated or secreted forms of the specific PRO polypeptide (e.g., an extracellular domain sequence), naturally-occurring variant forms (e.g., alternatively spliced forms) and naturally-occurring allelic variants of the polypeptide. In various

embodiments of the invention, the native sequence PRO polypeptides disclosed herein are mature or full-length native sequence polypeptides comprising the full-length amino acids sequences shown in the accompanying figures. Start and stop codons are shown in bold font and underlined in the figures. However, while the PRO polypeptide disclosed in the accompanying figures are shown to begin with methionine residues designated herein as amino acid position 1 in the figures, it is conceivable and possible that other methionine residues located either upstream or downstream from the amino acid position 1 in the figures may be employed as the starting amino acid residue for the PRO polypeptides.

[0197] The PRO polypeptide “extracellular domain” or “ECD” refers to a form of the PRO polypeptide which is essentially free of the transmembrane and cytoplasmic domains. Ordinarily, a PRO polypeptide ECD will have less than 1% of such transmembrane and/or cytoplasmic domains and preferably, will have less than 0.5% of such domains. It will be understood that any transmembrane domains identified for the PRO polypeptides of the present invention are identified pursuant to criteria routinely employed in the art for identifying that type of hydrophobic domain. The exact boundaries of a transmembrane domain may vary but most likely by no more than about 5 amino acids at either end of the domain as initially identified herein. Optionally, therefore, an extracellular domain of a PRO polypeptide may contain from about 5 or fewer amino acids on either side of the transmembrane domain/extracellular domain boundary as identified in the Examples or specification and such polypeptides, with or without the associated signal peptide, and nucleic acid encoding them, are contemplated by the present invention.

[0198] The approximate location of the “signal peptides” of the various PRO polypeptides disclosed herein are shown in the present specification and/or the accompanying figures. It is noted, however, that the C-terminal boundary of a signal peptide may vary, but most likely by no more than about 5 amino acids on either side of the signal peptide C-terminal boundary as initially identified herein, wherein the C-terminal boundary of the signal peptide may be identified pursuant to criteria routinely employed in the art for identifying that type of amino acid sequence element (e.g., Nielsen et al., Prot. Eng. 10:1-6 (1997) and von Heinje et al., Nucl. Acids. Res. 14:4683-4690 (1986)). Moreover, it is also recognized that, in some cases, cleavage of a signal sequence from a secreted polypeptide is not entirely uniform,

resulting in more than one secreted species. These mature polypeptides, where the signal peptide is cleaved within no more than about 5 amino acids on either side of the C-terminal boundary of the signal peptide as identified herein, and the polynucleotides encoding them, are contemplated by the present invention.

[0199] “PRO polypeptide variant” means an active PRO polypeptide as defined above or below having at least about 80% amino acid sequence identity with a full-length native sequence PRO polypeptide sequence as disclosed herein, a PRO polypeptide sequence lacking the signal peptide as disclosed herein, an extracellular domain of a PRO polypeptide, with or without the signal peptide, as disclosed herein or any other fragment of a full-length PRO polypeptide sequence as disclosed herein. Such PRO polypeptide variants include, for instance, PRO polypeptides wherein one or more amino acid residues are added, or deleted, at the N- or C-terminus of the full-length native amino acid sequence. Ordinarily, a PRO polypeptide variant will have at least about 80% amino acid sequence identity, alternatively at least about 81% amino acid sequence identity, alternatively at least about 82% amino acid sequence identity, alternatively at least about 83% amino acid sequence identity, alternatively at least about 84% amino acid sequence identity, alternatively at least about 85% amino acid sequence identity, alternatively at least about 86% amino acid sequence identity, alternatively at least about 87% amino acid sequence identity, alternatively at least about 88% amino acid sequence identity, alternatively at least about 89% amino acid sequence identity, alternatively at least about 90% amino acid sequence identity, alternatively at least about 91% amino acid sequence identity, alternatively at least about 92% amino acid sequence identity, alternatively at least about 93% amino acid sequence identity, alternatively at least about 94% amino acid sequence identity, alternatively at least about 95% amino acid sequence identity, alternatively at least about 96% amino acid sequence identity, alternatively at least about 97% amino acid sequence identity, alternatively at least about 98% amino acid sequence identity and alternatively at least about 99% amino acid sequence identity to a full-length native sequence PRO polypeptide sequence as disclosed herein, a PRO polypeptide sequence lacking the signal peptide as disclosed herein, an extracellular domain of a PRO polypeptide, with or without the signal peptide, as disclosed herein or any other specifically defined fragment of a full-length PRO polypeptide sequence as disclosed herein. Ordinarily, PRO variant polypeptides are at

least about 10 amino acids in length, alternatively at least about 20 amino acids in length, alternatively at least about 30 amino acids in length, alternatively at least about 40 amino acids in length, alternatively at least about 50 amino acids in length, alternatively at least about 60 amino acids in length, alternatively at least about 70 amino acids in length, alternatively at least about 80 amino acids in length, alternatively at least about 90 amino acids in length, alternatively at least about 100 amino acids in length, alternatively at least about 150 amino acids in length, alternatively at least about 200 amino acids in length, alternatively at least about 300 amino acids in length, or more.

[0200] “Percent (%) amino acid sequence identity” with respect to the PRO polypeptide sequences identified herein is defined as the percentage of amino acid residues in a candidate sequence that are identical with the amino acid residues in the specific PRO polypeptide sequence, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity, and not considering any conservative substitutions as part of the sequence identity. Alignment for purposes of determining percent amino acid sequence identity can be achieved in various ways that are within the skill in the art, for instance, using publicly available computer software such as BLAST, BLAST-2, ALIGN or Megalign (DNASTAR) software. Those skilled in the art can determine appropriate parameters for measuring alignment, including any algorithms needed to achieve maximal alignment over the full length of the sequences being compared. For purposes herein, however, % amino acid sequence identity values are generated using the sequence comparison computer program ALIGN-2, wherein the complete source code for the ALIGN-2 program is provided in Table 1 below. The ALIGN-2 sequence comparison computer program was authored by Genentech, Inc. and the source code shown in Table 1 below has been filed with user documentation in the U.S. Copyright Office, Washington D.C., 20559, where it is registered under U.S. Copyright Registration No. TXU510087. The ALIGN-2 program is publicly available through Genentech, Inc., South San Francisco, California or may be compiled from the source code provided in Table 1 below. The ALIGN-2 program should be compiled for use on a UNIX operating system, preferably digital UNIX V4.0D. All sequence comparison parameters are set by the ALIGN-2 program and do not vary.

[0201] In situations where ALIGN-2 is employed for amino acid sequence comparisons, the % amino acid sequence identity of a given amino acid sequence A to, with, or against a given amino acid sequence B (which can alternatively be phrased as a given amino acid sequence A that has or comprises a certain % amino acid sequence identity to, with, or against a given amino acid sequence B) is calculated as follows:

$$100 \text{ times the fraction } X/Y$$

where X is the number of amino acid residues scored as identical matches by the sequence alignment program ALIGN-2 in that program's alignment of A and B, and where Y is the total number of amino acid residues in B. It will be appreciated that where the length of amino acid sequence A is not equal to the length of amino acid sequence B, the % amino acid sequence identity of A to B will not equal the % amino acid sequence identity of B to A. As examples of % amino acid sequence identity calculations using this method, Tables 2 and 3 demonstrate how to calculate the % amino acid sequence identity of the amino acid sequence designated "Comparison Protein" to the amino acid sequence designated "PRO", wherein "PRO" represents the amino acid sequence of a hypothetical PRO polypeptide of interest, "Comparison Protein" represents the amino acid sequence of a polypeptide against which the "PRO" polypeptide of interest is being compared, and "X," "Y" and "Z" each represent different hypothetical amino acid residues.

[0202] Unless specifically stated otherwise, all % amino acid sequence identity values used herein are obtained as described in the immediately preceding paragraph using the ALIGN-2 computer program. However, % amino acid sequence identity values may also be obtained as described below by using the WU-BLAST-2 computer program (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Most of the WU-BLAST-2 search parameters are set to the default values. Those not set to default values, i.e., the adjustable parameters, are set with the following values: overlap span = 1, overlap fraction = 0.125, word threshold (T) = 11, and scoring matrix = BLOSUM62. When WU-BLAST-2 is employed, a % amino acid sequence identity value is determined by dividing (a) the number of matching identical amino acid residues between the amino acid sequence of the PRO polypeptide of interest having a

sequence derived from the native PRO polypeptide and the comparison amino acid sequence of interest (i.e., the sequence against which the PRO polypeptide of interest is being compared which may be a PRO variant polypeptide) as determined by WU-BLAST-2 by (b) the total number of amino acid residues of the PRO polypeptide of interest. For example, in the statement “a polypeptide comprising an the amino acid sequence A which has or having at least 80% amino acid sequence identity to the amino acid sequence B”, the amino acid sequence A is the comparison amino acid sequence of interest and the amino acid sequence B is the amino acid sequence of the PRO polypeptide of interest.

[0203] Percent amino acid sequence identity may also be determined using the sequence comparison program NCBI-BLAST2 (Altschul et al., Nucleic Acids Res. 25:3389-3402 (1997)). The NCBI-BLAST2 sequence comparison program may be downloaded from <http://www.ncbi.nlm.nih.gov> the National Institutes of Health website or otherwise obtained from the National Institute of Health, Bethesda, MD. NCBI-BLAST2 uses several search parameters, wherein all of those search parameters are set to default values including, for example, unmask = yes, strand = all, expected occurrences = 10, minimum low complexity length = 15/5, multi-pass e-value = 0.01, constant for multi-pass = 25, dropoff for final gapped alignment = 25 and scoring matrix = BLOSUM62.

[0204] In situations where NCBI-BLAST2 is employed for amino acid sequence comparisons, the % amino acid sequence identity of a given amino acid sequence A to, with, or against a given amino acid sequence B (which can alternatively be phrased as a given amino acid sequence A that has or comprises a certain % amino acid sequence identity to, with, or against a given amino acid sequence B) is calculated as follows:

$$100 \text{ times the fraction } X/Y$$

where X is the number of amino acid residues scored as identical matches by the sequence alignment program NCBI-BLAST2 in that program's alignment of A and B, and where Y is the total number of amino acid residues in B. It will be appreciated that where the length of amino acid sequence A is not equal to the length of amino acid sequence B, the % amino acid sequence identity of A to B will not equal the % amino acid sequence identity of B to A.

**[0205]** “PRO variant polynucleotide” or “PRO variant nucleic acid sequence” means a nucleic acid molecule which encodes an active PRO polypeptide as defined below and which has at least about 80% nucleic acid sequence identity with a nucleotide acid sequence encoding a full-length native sequence PRO polypeptide sequence as disclosed herein, a full-length native sequence PRO polypeptide sequence lacking the signal peptide as disclosed herein, an extracellular domain of a PRO polypeptide, with or without the signal peptide, as disclosed herein or any other fragment of a full-length PRO polypeptide sequence as disclosed herein. Ordinarily, a PRO variant polynucleotide will have at least about 80% nucleic acid sequence identity, alternatively at least about 81% nucleic acid sequence identity, alternatively at least about 82% nucleic acid sequence identity, alternatively at least about 83% nucleic acid sequence identity, alternatively at least about 84% nucleic acid sequence identity, alternatively at least about 85% nucleic acid sequence identity, alternatively at least about 86% nucleic acid sequence identity, alternatively at least about 87% nucleic acid sequence identity, alternatively at least about 88% nucleic acid sequence identity, alternatively at least about 89% nucleic acid sequence identity, alternatively at least about 90% nucleic acid sequence identity, alternatively at least about 91% nucleic acid sequence identity, alternatively at least about 92% nucleic acid sequence identity, alternatively at least about 93% nucleic acid sequence identity, alternatively at least about 94% nucleic acid sequence identity, alternatively at least about 95% nucleic acid sequence identity, alternatively at least about 96% nucleic acid sequence identity, alternatively at least about 97% nucleic acid sequence identity, alternatively at least about 98% nucleic acid sequence identity and alternatively at least about 99% nucleic acid sequence identity with a nucleic acid sequence encoding a full-length native sequence PRO polypeptide sequence as disclosed herein, a full-length native sequence PRO polypeptide sequence lacking the signal peptide as disclosed herein, an extracellular domain of a PRO polypeptide, with or without the signal sequence, as disclosed herein or any other fragment of a full-length PRO polypeptide sequence as disclosed herein. Variants do not encompass the native nucleotide sequence.

**[0206]** Ordinarily, PRO variant polynucleotides are at least about 30 nucleotides in length, alternatively at least about 60 nucleotides in length, alternatively at least about 90 nucleotides in length, alternatively at least about 120 nucleotides in length, alternatively at least about 150 nucleotides in length, alternatively at least about 180 nucleotides in length,



alternatively at least about 210 nucleotides in length, alternatively at least about 240 nucleotides in length, alternatively at least about 270 nucleotides in length, alternatively at least about 300 nucleotides in length, alternatively at least about 450 nucleotides in length, alternatively at least about 600 nucleotides in length, alternatively at least about 900 nucleotides in length, or more.

[0207] “Percent (%) nucleic acid sequence identity” with respect to PRO-encoding nucleic acid sequences identified herein is defined as the percentage of nucleotides in a candidate sequence that are identical with the nucleotides in the PRO nucleic acid sequence of interest, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity. Alignment for purposes of determining percent nucleic acid sequence identity can be achieved in various ways that are within the skill in the art, for instance, using publicly available computer software such as BLAST, BLAST-2, ALIGN or Megalign (DNASTAR) software. For purposes herein, however, % nucleic acid sequence identity values are generated using the sequence comparison computer program ALIGN-2, wherein the complete source code for the ALIGN-2 program is provided in Table 1 below. The ALIGN-2 sequence comparison computer program was authored by Genentech, Inc. and the source code shown in Table 1 below has been filed with user documentation in the U.S. Copyright Office, Washington D.C., 20559, where it is registered under U.S. Copyright Registration No. TXU510087. The ALIGN-2 program is publicly available through Genentech, Inc., South San Francisco, California or may be compiled from the source code provided in Table 1 below. The ALIGN-2 program should be compiled for use on a UNIX operating system, preferably digital UNIX V4.0D. All sequence comparison parameters are set by the ALIGN-2 program and do not vary.

[0208] In situations where ALIGN-2 is employed for nucleic acid sequence comparisons, the % nucleic acid sequence identity of a given nucleic acid sequence C to, with, or against a given nucleic acid sequence D (which can alternatively be phrased as a given nucleic acid sequence C that has or comprises a certain % nucleic acid sequence identity to, with, or against a given nucleic acid sequence D) is calculated as follows:

$$100 \text{ times the fraction } W/Z$$

where W is the number of nucleotides scored as identical matches by the sequence alignment program ALIGN-2 in that program's alignment of C and D, and where Z is the total number of nucleotides in D. It will be appreciated that where the length of nucleic acid sequence C is not equal to the length of nucleic acid sequence D, the % nucleic acid sequence identity of C to D will not equal the % nucleic acid sequence identity of D to C. As examples of % nucleic acid sequence identity calculations, Tables 4 and 5, demonstrate how to calculate the % nucleic acid sequence identity of the nucleic acid sequence designated "Comparison DNA" to the nucleic acid sequence designated "PRO-DNA", wherein "PRO-DNA" represents a hypothetical PRO-encoding nucleic acid sequence of interest, "Comparison DNA" represents the nucleotide sequence of a nucleic acid molecule against which the "PRO-DNA" nucleic acid molecule of interest is being compared, and "N", "L" and "V" each represent different hypothetical nucleotides.

[0209] Unless specifically stated otherwise, all % nucleic acid sequence identity values used herein are obtained as described in the immediately preceding paragraph using the ALIGN-2 computer program. However, % nucleic acid sequence identity values may also be obtained as described below by using the WU-BLAST-2 computer program (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Most of the WU-BLAST-2 search parameters are set to the default values. Those not set to default values, i.e., the adjustable parameters, are set with the following values: overlap span = 1, overlap fraction = 0.125, word threshold (T) = 11, and scoring matrix = BLOSUM62. When WU-BLAST-2 is employed, a % nucleic acid sequence identity value is determined by dividing (a) the number of matching identical nucleotides between the nucleic acid sequence of the PRO polypeptide-encoding nucleic acid molecule of interest having a sequence derived from the native sequence PRO polypeptide-encoding nucleic acid and the comparison nucleic acid molecule of interest (i.e., the sequence against which the PRO polypeptide-encoding nucleic acid molecule of interest is being compared which may be a variant PRO polynucleotide) as determined by WU-BLAST-2 by (b) the total number of nucleotides of the PRO polypeptide-encoding nucleic acid molecule of interest. For example, in the statement "an isolated nucleic acid molecule comprising a nucleic acid sequence A which has or having at least 80% nucleic acid sequence identity to the

nucleic acid sequence B”, the nucleic acid sequence A is the comparison nucleic acid molecule of interest and the nucleic acid sequence B is the nucleic acid sequence of the PRO polypeptide-encoding nucleic acid molecule of interest.

[0210] Percent nucleic acid sequence identity may also be determined using the sequence comparison program NCBI-BLAST2 (Altschul et al., Nucleic Acids Res. 25:3389-3402 (1997)). The NCBI-BLAST2 sequence comparison program may be downloaded from <http://www.ncbi.nlm.nih.gov> the National Institutes of Health website or otherwise obtained from the National Institute of Health, Bethesda, MD. NCBI-BLAST2 uses several search parameters, wherein all of those search parameters are set to default values including, for example, unmask = yes, strand = all, expected occurrences = 10, minimum low complexity length = 15/5, multi-pass e-value = 0.01, constant for multi-pass = 25, dropoff for final gapped alignment = 25 and scoring matrix = BLOSUM62.

[0211] In situations where NCBI-BLAST2 is employed for sequence comparisons, the % nucleic acid sequence identity of a given nucleic acid sequence C to, with, or against a given nucleic acid sequence D (which can alternatively be phrased as a given nucleic acid sequence C that has or comprises a certain % nucleic acid sequence identity to, with, or against a given nucleic acid sequence D) is calculated as follows:

$$100 \text{ times the fraction } W/Z$$

where W is the number of nucleotides scored as identical matches by the sequence alignment program NCBI-BLAST2 in that program’s alignment of C and D, and where Z is the total number of nucleotides in D. It will be appreciated that where the length of nucleic acid sequence C is not equal to the length of nucleic acid sequence D, the % nucleic acid sequence identity of C to D will not equal the % nucleic acid sequence identity of D to C.

[0212] In other embodiments, PRO variant polynucleotides are nucleic acid molecules that encode an active PRO polypeptide and which are capable of hybridizing, preferably under stringent hybridization and wash conditions, to nucleotide sequences encoding a full-length PRO polypeptide as disclosed herein. PRO variant polypeptides may be those that are encoded by a PRO variant polynucleotide.

[0213] “Isolated,” when used to describe the various polypeptides disclosed herein, means polypeptide that has been identified and separated and/or recovered from a component of its natural environment. Contaminant components of its natural environment are materials that would typically interfere with diagnostic or therapeutic uses for the polypeptide, and may include enzymes, hormones, and other proteinaceous or non-proteinaceous solutes. In preferred embodiments, the polypeptide will be purified (1) to a degree sufficient to obtain at least 15 residues of N-terminal or internal amino acid sequence by use of a spinning cup sequenator, or (2) to homogeneity by SDS-PAGE under non-reducing or reducing conditions using Coomassie blue or, preferably, silver stain. Isolated polypeptide includes polypeptide *in situ* within recombinant cells, since at least one component of the PRO polypeptide natural environment will not be present. Ordinarily, however, isolated polypeptide will be prepared by at least one purification step.

[0214] An “isolated” PRO polypeptide-encoding nucleic acid or other polypeptide-encoding nucleic acid is a nucleic acid molecule that is identified and separated from at least one contaminant nucleic acid molecule with which it is ordinarily associated in the natural source of the polypeptide-encoding nucleic acid. An isolated polypeptide-encoding nucleic acid molecule is other than in the form or setting in which it is found in nature. Isolated polypeptide-encoding nucleic acid molecules therefore are distinguished from the specific polypeptide-encoding nucleic acid molecule as it exists in natural cells. However, an isolated polypeptide-encoding nucleic acid molecule includes polypeptide-encoding nucleic acid molecules contained in cells that ordinarily express the polypeptide where, for example, the nucleic acid molecule is in a chromosomal location different from that of natural cells.

[0215] The term “control sequences” refers to DNA sequences necessary for the expression of an operably linked coding sequence in a particular host organism. The control sequences that are suitable for prokaryotes, for example, include a promoter, optionally an operator sequence, and a ribosome binding site. Eukaryotic cells are known to utilize promoters, polyadenylation signals, and enhancers.

[0216] Nucleic acid is “operably linked” when it is placed into a functional relationship with another nucleic acid sequence. For example, DNA for a presequence or secretory leader is operably linked to DNA for a polypeptide if it is expressed as a preprotein

that participates in the secretion of the polypeptide; a promoter or enhancer is operably linked to a coding sequence if it affects the transcription of the sequence; or a ribosome binding site is operably linked to a coding sequence if it is positioned so as to facilitate translation. Generally, “operably linked” means that the DNA sequences being linked are contiguous, and, in the case of a secretory leader, contiguous and in reading phase. However, enhancers do not have to be contiguous. Linking is accomplished by ligation at convenient restriction sites. If such sites do not exist, the synthetic oligonucleotide adaptors or linkers are used in accordance with conventional practice.

[0217] The term “antibody” is used in the broadest sense and specifically covers, for example, single anti-PRO monoclonal antibodies (including agonist, antagonist, and neutralizing antibodies), anti-PRO antibody compositions with polyepitopic specificity, single chain anti-PRO antibodies, and fragments of anti-PRO antibodies (see below). The term “monoclonal antibody” as used herein refers to an antibody obtained from a population of substantially homogeneous antibodies, i.e., the individual antibodies comprising the population are identical except for possible naturally-occurring mutations that may be present in minor amounts.

[0218] “Stringency” of hybridization reactions is readily determinable by one of ordinary skill in the art, and generally is an empirical calculation dependent upon probe length, washing temperature, and salt concentration. In general, longer probes require higher temperatures for proper annealing, while shorter probes need lower temperatures. Hybridization generally depends on the ability of denatured DNA to reanneal when complementary strands are present in an environment below their melting temperature. The higher the degree of desired homology between the probe and hybridizable sequence, the higher the relative temperature which can be used. As a result, it follows that higher relative temperatures would tend to make the reaction conditions more stringent, while lower temperatures less so. For additional details and explanation of stringency of hybridization reactions, see Ausubel et al., Current Protocols in Molecular Biology, Wiley Interscience Publishers, (1995).

[0219] “Stringent conditions” or “high stringency conditions”, as defined herein, may be identified by those that: (1) employ low ionic strength and high temperature for

washing, for example 0.015 M sodium chloride/0.0015 M sodium citrate/0.1% sodium dodecyl sulfate at 50°C; (2) employ during hybridization a denaturing agent, such as formamide, for example, 50% (v/v) formamide with 0.1% bovine serum albumin/0.1% Ficoll/0.1% polyvinylpyrrolidone/50mM sodium phosphate buffer at pH 6.5 with 750 mM sodium chloride, 75 mM sodium citrate at 42°C; or (3) employ 50% formamide, 5 x SSC (0.75 M NaCl, 0.075 M sodium citrate), 50 mM sodium phosphate (pH 6.8), 0.1% sodium pyrophosphate, 5 x Denhardt's solution, sonicated salmon sperm DNA (50 µg/ml), 0.1% SDS, and 10% dextran sulfate at 42°C, with washes at 42°C in 0.2 x SSC (sodium chloride/sodium citrate) and 50% formamide at 55°C, followed by a high-stringency wash consisting of 0.1 x SSC containing EDTA at 55°C.

[0220] “Moderately stringent conditions” may be identified as described by Sambrook et al., Molecular Cloning: A Laboratory Manual, New York: Cold Spring Harbor Press, 1989, and include the use of washing solution and hybridization conditions (e.g., temperature, ionic strength and %SDS) less stringent than those described above. An example of moderately stringent conditions is overnight incubation at 37°C in a solution comprising: 20% formamide, 5 x SSC (150 mM NaCl, 15 mM trisodium citrate), 50 mM sodium phosphate (pH 7.6), 5 x Denhardt's solution, 10% dextran sulfate, and 20 mg/ml denatured sheared salmon sperm DNA, followed by washing the filters in 1 x SSC at about 37-50°C. The skilled artisan will recognize how to adjust the temperature, ionic strength, etc. as necessary to accommodate factors such as probe length and the like.

[0221] The term “epitope tagged” when used herein refers to a chimeric polypeptide comprising a PRO polypeptide fused to a “tag polypeptide”. The tag polypeptide has enough residues to provide an epitope against which an antibody can be made, yet is short enough such that it does not interfere with activity of the polypeptide to which it is fused. The tag polypeptide preferably also is fairly unique so that the antibody does not substantially cross-react with other epitopes. Suitable tag polypeptides generally have at least six amino acid residues and usually between about 8 and 50 amino acid residues (preferably, between about 10 and 20 amino acid residues).

[0222] As used herein, the term “immunoadhesin” designates antibody-like molecules which combine the binding specificity of a heterologous protein (an “adhesin”) with

the effector functions of immunoglobulin constant domains. Structurally, the immunoadhesins comprise a fusion of an amino acid sequence with the desired binding specificity which is other than the antigen recognition and binding site of an antibody (i.e., is “heterologous”), and an immunoglobulin constant domain sequence. The adhesin part of an immunoadhesin molecule typically is a contiguous amino acid sequence comprising at least the binding site of a receptor or a ligand. The immunoglobulin constant domain sequence in the immunoadhesin may be obtained from any immunoglobulin, such as IgG-1, IgG-2, IgG-3, or IgG-4 subtypes, IgA (including IgA-1 and IgA-2), IgE, IgD or IgM.

**[0223]** “Active” or “activity” for the purposes herein refers to form(s) of a PRO polypeptide which retain a biological and/or an immunological activity of native or naturally-occurring PRO, wherein “biological” activity refers to a biological function (either inhibitory or stimulatory) caused by a native or naturally-occurring PRO other than the ability to induce the production of an antibody against an antigenic epitope possessed by a native or naturally-occurring PRO and an “immunological” activity refers to the ability to induce the production of an antibody against an antigenic epitope possessed by a native or naturally-occurring PRO.

**[0224]** The term “antagonist” is used in the broadest sense, and includes any molecule that partially or fully blocks, inhibits, or neutralizes a biological activity of a native PRO polypeptide disclosed herein. In a similar manner, the term “agonist” is used in the broadest sense and includes any molecule that mimics a biological activity of a native PRO polypeptide disclosed herein. Suitable agonist or antagonist molecules specifically include agonist or antagonist antibodies or antibody fragments, fragments or amino acid sequence variants of native PRO polypeptides, peptides, antisense oligonucleotides, small organic molecules, etc. Methods for identifying agonists or antagonists of a PRO polypeptide may comprise contacting a PRO polypeptide with a candidate agonist or antagonist molecule and measuring a detectable change in one or more biological activities normally associated with the PRO polypeptide.

**[0225]** “Treatment” refers to both therapeutic treatment and prophylactic or preventative measures, wherein the object is to prevent or slow down (lessen) the targeted pathologic condition or disorder. Those in need of treatment include those already with the

disorder as well as those prone to have the disorder or those in whom the disorder is to be prevented.

[0226] “Chronic” administration refers to administration of the agent(s) in a continuous mode as opposed to an acute mode, so as to maintain the initial therapeutic effect (activity) for an extended period of time. “Intermittent” administration is treatment that is not consecutively done without interruption, but rather is cyclic in nature.

[0227] “Mammal” for purposes of treatment refers to any animal classified as a mammal, including humans, domestic and farm animals, and zoo, sports, or pet animals, such as dogs, cats, cattle, horses, sheep, pigs, goats, rabbits, etc. Preferably, the mammal is human.

[0228] Administration “in combination with” one or more further therapeutic agents includes simultaneous (concurrent) and consecutive administration in any order.

[0229] “Carriers” as used herein include pharmaceutically acceptable carriers, excipients, or stabilizers which are nontoxic to the cell or mammal being exposed thereto at the dosages and concentrations employed. Often the physiologically acceptable carrier is an aqueous pH buffered solution. Examples of physiologically acceptable carriers include buffers such as phosphate, citrate, and other organic acids; antioxidants including ascorbic acid; low molecular weight (less than about 10 residues) polypeptide; proteins, such as serum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone; amino acids such as glycine, glutamine, asparagine, arginine or lysine; monosaccharides, disaccharides, and other carbohydrates including glucose, mannose, or dextrans; chelating agents such as EDTA; sugar alcohols such as mannitol or sorbitol; salt-forming counterions such as sodium; and/or nonionic surfactants such as TWEEN™, polyethylene glycol (PEG), and PLURONICS™.

[0230] “Antibody fragments” comprise a portion of an intact antibody, preferably the antigen binding or variable region of the intact antibody. Examples of antibody fragments include Fab, Fab’, F(ab’)<sub>2</sub>, and Fv fragments; diabodies; linear antibodies (Zapata et al., Protein Eng. 8(10): 1057-1062 [1995]); single-chain antibody molecules; and multispecific antibodies formed from antibody fragments.

[0231] Papain digestion of antibodies produces two identical antigen-binding fragments, called “Fab” fragments, each with a single antigen-binding site, and a residual “Fc” fragment, a designation reflecting the ability to crystallize readily. Pepsin treatment yields an



F(ab')<sub>2</sub> fragment that has two antigen-combining sites and is still capable of cross-linking antigen.

[0232] “Fv” is the minimum antibody fragment which contains a complete antigen-recognition and -binding site. This region consists of a dimer of one heavy- and one light-chain variable domain in tight, non-covalent association. It is in this configuration that the three CDRs of each variable domain interact to define an antigen-binding site on the surface of the V<sub>H</sub>-V<sub>L</sub> dimer. Collectively, the six CDRs confer antigen-binding specificity to the antibody. However, even a single variable domain (or half of an Fv comprising only three CDRs specific for an antigen) has the ability to recognize and bind antigen, although at a lower affinity than the entire binding site.

[0233] The Fab fragment also contains the constant domain of the light chain and the first constant domain (CH1) of the heavy chain. Fab fragments differ from Fab' fragments by the addition of a few residues at the carboxy terminus of the heavy chain CH1 domain including one or more cysteines from the antibody hinge region. Fab'-SH is the designation herein for Fab' in which the cysteine residue(s) of the constant domains bear a free thiol group. F(ab')<sub>2</sub> antibody fragments originally were produced as pairs of Fab' fragments which have hinge cysteines between them. Other chemical couplings of antibody fragments are also known.

[0234] The “light chains” of antibodies (immunoglobulins) from any vertebrate species can be assigned to one of two clearly distinct types, called kappa and lambda, based on the amino acid sequences of their constant domains.

[0235] Depending on the amino acid sequence of the constant domain of their heavy chains, immunoglobulins can be assigned to different classes. There are five major classes of immunoglobulins: IgA, IgD, IgE, IgG, and IgM, and several of these may be further divided into subclasses (isotypes), e.g., IgG1, IgG2, IgG3, IgG4, IgA, and IgA2.

[0236] “Single-chain Fv” or “sFv” antibody fragments comprise the V<sub>H</sub> and V<sub>L</sub> domains of antibody, wherein these domains are present in a single polypeptide chain. Preferably, the Fv polypeptide further comprises a polypeptide linker between the V<sub>H</sub> and V<sub>L</sub> domains which enables the sFv to form the desired structure for antigen binding. For a review

of sFv, see Pluckthun in The Pharmacology of Monoclonal Antibodies, vol. 113, Rosenberg and Moore eds., Springer-Verlag, New York, pp. 269-315 (1994).

[0237] The term “diabodies” refers to small antibody fragments with two antigen-binding sites, which fragments comprise a heavy-chain variable domain ( $V_H$ ) connected to a light-chain variable domain ( $V_L$ ) in the same polypeptide chain ( $V_H$ - $V_L$ ). By using a linker that is too short to allow pairing between the two domains on the same chain, the domains are forced to pair with the complementary domains of another chain and create two antigen-binding sites. Diabodies are described more fully in, for example, EP 404,097; WO 93/11161; and Hollinger et al., Proc. Natl. Acad. Sci. USA, 90:6444-6448 (1993).

[0238] An “isolated” antibody is one which has been identified and separated and/or recovered from a component of its natural environment. Contaminant components of its natural environment are materials which would interfere with diagnostic or therapeutic uses for the antibody, and may include enzymes, hormones, and other proteinaceous or nonproteinaceous solutes. In preferred embodiments, the antibody will be purified (1) to greater than 95% by weight of antibody as determined by the Lowry method, and most preferably more than 99% by weight, (2) to a degree sufficient to obtain at least 15 residues of N-terminal or internal amino acid sequence by use of a spinning cup sequenator, or (3) to homogeneity by SDS-PAGE under reducing or nonreducing conditions using Coomassie blue or, preferably, silver stain. Isolated antibody includes the antibody in situ within recombinant cells since at least one component of the antibody’s natural environment will not be present. Ordinarily, however, isolated antibody will be prepared by at least one purification step.

[0239] An antibody that “specifically binds to” or is “specific for” a particular polypeptide or an epitope on a particular polypeptide is one that binds to that particular polypeptide or epitope on a particular polypeptide without substantially binding to any other polypeptide or polypeptide epitope.

[0240] The word “label” when used herein refers to a detectable compound or composition which is conjugated directly or indirectly to the antibody so as to generate a “labeled” antibody. The label may be detectable by itself (e.g. radioisotope labels or fluorescent labels) or, in the case of an enzymatic label, may catalyze chemical alteration of a substrate compound or composition which is detectable.

**[0241]** By “solid phase” is meant a non-aqueous matrix to which the antibody of the present invention can adhere. Examples of solid phases encompassed herein include those formed partially or entirely of glass (e.g., controlled pore glass), polysaccharides (e.g., agarose), polyacrylamides, polystyrene, polyvinyl alcohol and silicones. In certain embodiments, depending on the context, the solid phase can comprise the well of an assay plate; in others it is a purification column (e.g., an affinity chromatography column). This term also includes a discontinuous solid phase of discrete particles, such as those described in U.S. Patent No. 4,275,149.

**[0242]** A “liposome” is a small vesicle composed of various types of lipids, phospholipids and/or surfactant which is useful for delivery of a drug (such as a PRO polypeptide or antibody thereto) to a mammal. The components of the liposome are commonly arranged in a bilayer formation, similar to the lipid arrangement of biological membranes.

**[0243]** A “small molecule” is defined herein to have a molecular weight below about 500 Daltons.

**Table 1**

```

/*
 *
 * C-C increased from 12 to 15
 * Z is average of EQ
 * B is average of ND
 * match with stop is _M; stop-stop = 0; J (joker) match = 0
 */
#define _M      -8      /* value of a match with a stop */

int  _day[26][26] = {
/*   A B C D E F G H I J K L M N O P Q R S T U V W X Y Z */
/* A */ { 2, 0, -2, 0, 0, -4, 1, -1, -1, 0, -1, -2, -1, 0, _M, 1, 0, -2, 1, 1, 0, 0, -6, 0, -3, 0},
/* B */ { 0, 3, -4, 3, 2, -5, 0, 1, -2, 0, 0, -3, -2, 2, _M, -1, 1, 0, 0, 0, 0, -2, -5, 0, -3, 1},
/* C */ {-2, -4, 15, -5, -5, -4, -3, -3, -2, 0, -5, -6, -5, -4, _M, -3, -5, -4, 0, -2, 0, -2, -8, 0, 0, -5},
/* D */ { 0, 3, -5, 4, 3, -6, 1, 1, -2, 0, 0, -4, -3, 2, _M, -1, 2, -1, 0, 0, 0, -2, -7, 0, -4, 2},
/* E */ { 0, 2, -5, 3, 4, -5, 0, 1, -2, 0, 0, -3, -2, 1, _M, -1, 2, -1, 0, 0, 0, -2, -7, 0, -4, 3},
/* F */ {-4, -5, -4, -6, -5, 9, -5, -2, 1, 0, -5, 2, 0, -4, _M, -5, -5, -4, -3, -3, 0, -1, 0, 0, 7, -5},
/* G */ { 1, 0, -3, 1, 0, -5, 5, -2, -3, 0, -2, -4, -3, 0, _M, -1, -1, -3, 1, 0, 0, -1, -7, 0, -5, 0},
/* H */ {-1, 1, -3, 1, 1, -2, -2, 6, -2, 0, 0, -2, -2, 2, _M, 0, 3, 2, -1, -1, 0, -2, -3, 0, 0, 2},
/* I */ {-1, -2, -2, -2, -2, 1, -3, -2, 5, 0, -2, 2, 2, -2, _M, -2, -2, -2, -1, 0, 0, 4, -5, 0, -1, -2},
/* J */ { 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, _M, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0},
/* K */ {-1, 0, -5, 0, 0, -5, -2, 0, -2, 0, 5, -3, 0, 1, _M, -1, 1, 3, 0, 0, 0, -2, -3, 0, -4, 0},
/* L */ {-2, -3, -6, -4, -3, 2, -4, -2, 2, 0, -3, 6, 4, -3, _M, -3, -2, -3, -3, -1, 0, 2, -2, 0, -1, -2},
/* M */ {-1, -2, -5, -3, -2, 0, -3, -2, 2, 0, 0, 4, 6, -2, _M, -2, -1, 0, -2, -1, 0, 2, -4, 0, -2, -1},
/* N */ { 0, 2, -4, 2, 1, -4, 0, 2, -2, 0, 1, -3, -2, 2, _M, -1, 1, 0, 1, 0, 0, -2, -4, 0, -2, 1},
/* O */ { _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, 0, _M, _M, _M, _M, _M, _M, _M, _M, _M},
/* P */ { 1, -1, -3, -1, -1, -5, -1, 0, -2, 0, -1, -3, -2, -1, _M, 6, 0, 0, 1, 0, 0, -1, -6, 0, -5, 0},
/* Q */ { 0, 1, -5, 2, 2, -5, -1, 3, -2, 0, 1, -2, -1, 1, _M, 0, 4, 1, -1, -1, 0, -2, -5, 0, -4, 3},
/* R */ {-2, 0, -4, -1, -1, -4, -3, 2, -2, 0, 3, -3, 0, 0, _M, 0, 1, 6, 0, -1, 0, -2, 2, 0, -4, 0},
/* S */ { 1, 0, 0, 0, 0, -3, 1, -1, -1, 0, 0, -3, -2, 1, _M, 1, -1, 0, 2, 1, 0, -1, -2, 0, -3, 0},
/* T */ { 1, 0, -2, 0, 0, -3, 0, -1, 0, 0, 0, -1, -1, 0, _M, 0, -1, -1, 1, 3, 0, 0, -5, 0, -3, 0},
/* U */ { 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, _M, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0},
/* V */ { 0, -2, -2, -2, -2, -1, -1, -2, 4, 0, -2, 2, 2, -2, _M, -1, -2, -2, -1, 0, 0, 4, -6, 0, -2, -2},
/* W */ {-6, -5, -8, -7, -7, 0, -7, -3, -5, 0, -3, -2, -4, -4, _M, -6, -5, 2, -2, -5, 0, -6, 17, 0, 0, -6},
/* X */ { 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, _M, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0},
/* Y */ {-3, -3, 0, -4, -4, 7, -5, 0, -1, 0, -4, -1, -2, -2, _M, -5, -4, -4, -3, -3, 0, -2, 0, 0, 10, -4},
/* Z */ { 0, 1, -5, 2, 3, -5, 0, 2, -2, 0, 0, -2, -1, 1, _M, 0, 3, 0, 0, 0, 0, -2, -6, 0, -4, 4}
};

```

**Table 1 (cont')**

```
/*
*/
#include <stdio.h>
#include <ctype.h>

#define MAXJMP      16      /* max jumps in a diag */
#define MAXGAP      24      /* don't continue to penalize gaps larger than this */
#define JMPS        1024    /* max jmps in an path */
#define MX          4       /* save if there's at least MX-1 bases since last jmp */

#define DMAT         3      /* value of matching bases */
#define DMIS         0      /* penalty for mismatched bases */
#define DINS0        8      /* penalty for a gap */
#define DINS1        1      /* penalty per base */
#define PINS0        8      /* penalty for a gap */
#define PINS1        4      /* penalty per residue */

struct jmp {
    short          n[MAXJMP];    /* size of jmp (neg for dely) */
    unsigned short x[MAXJMP];    /* base no. of jmp in seq x */
};
/* limits seq to 2^16 -1 */

struct diag {
    int            score;        /* score at last jmp */
    long           offset;       /* offset of prev block */
    short          ijmp;         /* current jmp index */
    struct jmp      jp;          /* list of jmps */
};

struct path {
    int            spc;          /* number of leading spaces */
    short          n[JMPS];      /* size of jmp (gap) */
    int            x[JMPS];      /* loc of jmp (last elem before gap) */
};

char              *ofile;        /* output file name */
char              *namex[2];     /* seq names: getseqs() */
char              *prog;         /* prog name for err msgs */
char              *seqx[2];      /* seqs: getseqs() */
int               dmax;          /* best diag: nw() */
int               dmax0;         /* final diag */
int               dna;           /* set if dna: main() */
int               endgaps;       /* set if penalizing end gaps */
int               gapx, gapy;     /* total gaps in seqs */
int               len0, len1;     /* seq lens */
int               ngapx, ngapy;   /* total size of gaps */
int               smax;          /* max score: nw() */
int               *xbm;          /* bitmap for matching */
long              offset;        /* current offset in jmp file */
struct            diag           /* holds diagonals */
struct            path           /* holds path for seqs */
    *dx;
    pp[2];

char              *calloc(), *malloc(), *index(), *strcpy();
char              *getseq(), *g_calloc();
```

## Table 1 (cont')

```

/* Needleman-Wunsch alignment program
*
* usage: progs file1 file2
* where file1 and file2 are two dna or two protein sequences.
* The sequences can be in upper- or lower-case and may contain ambiguity
* Any lines beginning with ';', '>' or '<' are ignored
* Max file length is 65535 (limited by unsigned short x in the jmp struct)
* A sequence with 1/3 or more of its elements ACGTU is assumed to be DNA
* Output is in the file "align.out"
*
* The program may create a tmp file in /tmp to hold info about traceback.
* Original version developed under BSD 4.3 on a vax 8650
*/
#include "nw.h"
#include "day.h"

static  _dbval[26] = {
    1,14,2,13,0,0,4,11,0,0,12,0,3,15,0,0,0,5,6,8,8,7,9,0,10,0
};

static  _pbval[26] = {
    1, 2[(1<<('D'-'A'))](1<<('N'-'A')), 4, 8, 16, 32, 64,
    128, 256, 0xFFFFFFFF, 1<<10, 1<<11, 1<<12, 1<<13, 1<<14,
    1<<15, 1<<16, 1<<17, 1<<18, 1<<19, 1<<20, 1<<21, 1<<22,
    1<<23, 1<<24, 1<<25[(1<<('E'-'A'))](1<<('Q'-'A'))
};

main(ac, av)
    int    ac;
    char    *av[ ];
{
    prog = av[0];
    if (ac != 3) {
        fprintf(stderr, "usage: %s file1 file2\n", prog);
        fprintf(stderr, "where file1 and file2 are two dna or two protein sequences.\n");
        fprintf(stderr, "The sequences can be in upper- or lower-case\n");
        fprintf(stderr, "Any lines beginning with ';' or '<' are ignored\n");
        fprintf(stderr, "Output is in the file \"align.out\"\n");
        exit(1);
    }
    namex[0] = av[1];
    namex[1] = av[2];
    seqx[0] = getseq(namex[0], &len0);
    seqx[1] = getseq(namex[1], &len1);
    xbm = (dna)? _dbval : _pbval;

    endgaps = 0;                /* 1 to penalize endgaps */
    ofile = "align.out";        /* output file */

    nw();                /* fill in the matrix, get the possible jmps */
    readjmps();          /* get the actual jmps */
    print();              /* print stats, alignment */

    cleanup(0);          /* unlink any tmp files */
}

```

**main**

## Table 1 (cont')

```

/* do the alignment, return best score: main()
 * dna: values in Fitch and Smith, PNAS, 80, 1382-1386, 1983
 * pro: PAM 250 values
 * When scores are equal, we prefer mismatches to any gap, prefer
 * a new gap to extending an ongoing gap, and prefer a gap in seqx
 * to a gap in seq y.
 */

```

```

nw()
{

```

nw

```

    char      *px, *py;          /* seqs and ptrs */
    int        *ndely, *dely;     /* keep track of dely */
    int        ndelx, delx;       /* keep track of delx */
    int        *tmp;              /* for swapping row0, row1 */
    int        mis;               /* score for each type */
    int        ins0, ins1;        /* insertion penalties */
    register   id;                /* diagonal index */
    register   ij;                /* jmp index */
    register   *col0, *col1;      /* score for curr, last row */
    register   xx, yy;            /* index into seqs */

```

```

    dx = (struct diag *)g_calloc("to get diags", len0+len1+1, sizeof(struct diag));

```

```

    ndely = (int *)g_calloc("to get ndely", len1+1, sizeof(int));
    dely = (int *)g_calloc("to get dely", len1+1, sizeof(int));
    col0 = (int *)g_calloc("to get col0", len1+1, sizeof(int));
    col1 = (int *)g_calloc("to get col1", len1+1, sizeof(int));
    ins0 = (dna)? DINS0 : PINS0;
    ins1 = (dna)? DINS1 : PINS1;

```

```

    smax = -10000;
    if (endgaps) {
        for (col0[0] = dely[0] = -ins0, yy = 1; yy <= len1; yy++) {
            col0[yy] = dely[yy] = col0[yy-1] - ins1;
            ndely[yy] = yy;
        }
        col0[0] = 0;          /* Waterman Bull Math Biol 84 */
    }
    else
        for (yy = 1; yy <= len1; yy++)
            dely[yy] = -ins0;

```

```

    /* fill in match matrix
    */

```

```

    for (px = seqx[0], xx = 1; xx <= len0; px++, xx++) {
        /* initialize first entry in col
        */
        if (endgaps) {
            if (xx == 1)
                col1[0] = delx = -(ins0+ins1);
            else
                col1[0] = delx = col0[0] - ins1;
            ndelx = xx;
        }
        else {
            col1[0] = 0;
            delx = -ins0;
            ndelx = 0;
        }
    }

```

### Table 1 (cont')

...nw

```
for (py = seqx[1], yy = 1; yy <= len1; py++, yy++) {
    mis = col0[yy-1];
    if (dna)
        mis += (xbm[*px-'A']&xbm[*py-'A'])? DMAT : DMIS;
    else
        mis += _day[*px-'A'][*py-'A'];

    /* update penalty for del in x seq;
     * favor new del over ongong del
     * ignore MAXGAP if weighting endgaps
     */
    if (endgaps || ndely[yy] < MAXGAP) {
        if (col0[yy] - ins0 >= dely[yy]) {
            dely[yy] = col0[yy] - (ins0+ins1);
            ndely[yy] = 1;
        } else {
            dely[yy] -= ins1;
            ndely[yy]++;
        }
    } else {
        if (col0[yy] - (ins0+ins1) >= dely[yy]) {
            dely[yy] = col0[yy] - (ins0+ins1);
            ndely[yy] = 1;
        } else
            ndely[yy]++;
    }

    /* update penalty for del in y seq;
     * favor new del over ongong del
     */
    if (endgaps || ndelx < MAXGAP) {
        if (col1[yy-1] - ins0 >= delx) {
            delx = col1[yy-1] - (ins0+ins1);
            ndelx = 1;
        } else {
            delx -= ins1;
            ndelx++;
        }
    } else {
        if (col1[yy-1] - (ins0+ins1) >= delx) {
            delx = col1[yy-1] - (ins0+ins1);
            ndelx = 1;
        } else
            ndelx++;
    }

    /* pick the maximum score; we're favoring
     * mis over any del and delx over dely
     */
}
```



**Table 1 (cont')**

...nw

```
id = xx - yy + len1 - 1;
if (mis >= delx && mis >= dely[yy])
    col1[yy] = mis;
else if (delx >= dely[yy]) {
    col1[yy] = delx;
    ij = dx[id].ijmp;
    if (dx[id].jp.n[0] && (!dna || (ndelx >= MAXJMP
    && xx > dx[id].jp.x[ij]+MX) || mis > dx[id].score+DINS0)) {
        dx[id].ijmp++;
        if (++ij >= MAXJMP) {
            writeimps(id);
            ij = dx[id].ijmp = 0;
            dx[id].offset = offset;
            offset += sizeof(struct jmp) + sizeof(offset);
        }
    }
    dx[id].jp.n[ij] = ndelx;
    dx[id].jp.x[ij] = xx;
    dx[id].score = delx;
}
else {
    col1[yy] = dely[yy];
    ij = dx[id].ijmp;
    if (dx[id].jp.n[0] && (!dna || (ndely[yy] >= MAXJMP
    && xx > dx[id].jp.x[ij]+MX) || mis > dx[id].score+DINS0)) {
        dx[id].ijmp++;
        if (++ij >= MAXJMP) {
            writeimps(id);
            ij = dx[id].ijmp = 0;
            dx[id].offset = offset;
            offset += sizeof(struct jmp) + sizeof(offset);
        }
    }
    dx[id].jp.n[ij] = -ndely[yy];
    dx[id].jp.x[ij] = xx;
    dx[id].score = dely[yy];
}
if (xx == len0 && yy < len1) {
    /* last col
    */
    if (endgaps)
        col1[yy] -= ins0+ins1*(len1-yy);
    if (col1[yy] > smax) {
        smax = col1[yy];
        dmax = id;
    }
}
}
if (endgaps && xx < len0)
    col1[yy-1] -= ins0+ins1*(len0-xx);
if (col1[yy-1] > smax) {
    smax = col1[yy-1];
    dmax = id;
}
}
tmp = col0; col0 = col1; col1 = tmp;
}
(void) free((char *)ndely);
(void) free((char *)dely);
```

**Table 1 (cont')**

```
(void) free((char *)col0);
(void) free((char *)col1);
}

/*
 *
 * print() – only routine visible outside this module
 *
 * static:
 * getmat() – trace back best path, count matches: print()
 * pr_align() – print alignment of described in array p[ ]: print()
 * dumpblock() – dump a block of lines with numbers, stars: pr_align()
 * nums() – put out a number line: dumpblock()
 * putline() – put out a line (name, [num], seq, [num]): dumpblock()
 * stars() – put a line of stars: dumpblock()
 * stripname() – strip any path and prefix from a seqname
 */

#include "nw.h"

#define SPC      3
#define P_LINE  256    /* maximum output line */
#define P_SPC    3      /* space between name or num and seq */

extern _day[26][26];
int      olen;          /* set output line length */
FILE     *fx;           /* output file */

print()
{
    int      lx, ly, firstgap, lastgap;    /* overlap */

    if ((fx = fopen(ofile, "w")) == 0) {
        fprintf(stderr, "%s: can't write %s\n", prog, ofile);
        cleanup(1);
    }
    fprintf(fx, "<first sequence: %s (length = %d)\n", namex[0], len0);
    fprintf(fx, "<second sequence: %s (length = %d)\n", namex[1], len1);
    olen = 60;
    lx = len0;
    ly = len1;
    firstgap = lastgap = 0;
    if (dmax < len1 - 1) { /* leading gap in x */
        pp[0].spc = firstgap = len1 - dmax - 1;
        ly -= pp[0].spc;
    }
    else if (dmax > len1 - 1) { /* leading gap in y */
        pp[1].spc = firstgap = dmax - (len1 - 1);
        lx -= pp[1].spc;
    }
    if (dmax0 < len0 - 1) { /* trailing gap in x */
        lastgap = len0 - dmax0 - 1;
        lx -= lastgap;
    }
    else if (dmax0 > len0 - 1) { /* trailing gap in y */
        lastgap = dmax0 - (len0 - 1);
        ly -= lastgap;
    }
    getmat(lx, ly, firstgap, lastgap);
    pr_align();
}
```

**print**

## Table 1 (cont')

```

/*
 * trace back the best path, count matches
 */
static
getmat(lx, ly, firstgap, lastgap)
    int      lx, ly;          /* "core" (minus endgaps) */
    int      firstgap, lastgap; /* leading trailing overlap */
{
    int      nm, i0, i1, siz0, siz1;
    char     outx[32];
    double   pct;
    register n0, n1;
    register char *p0, *p1;

    /* get total matches, score
     */
    i0 = i1 = siz0 = siz1 = 0;
    p0 = seqx[0] + pp[1].spc;
    p1 = seqx[1] + pp[0].spc;
    n0 = pp[1].spc + 1;
    n1 = pp[0].spc + 1;

    nm = 0;
    while ( *p0 && *p1 ) {
        if (siz0) {
            p1++;
            n1++;
            siz0--;
        }
        else if (siz1) {
            p0++;
            n0++;
            siz1--;
        }
        else {
            if (xbm[*p0-'A']&xbm[*p1-'A'])
                nm++;
            if (n0++ == pp[0].x[i0])
                siz0 = pp[0].n[i0++];
            if (n1++ == pp[1].x[i1])
                siz1 = pp[1].n[i1++];
            p0++;
            p1++;
        }
    }

    /* pct homology:
     * if penalizing endgaps, base is the shorter seq
     * else, knock off overhangs and take shorter core
     */
    if (endgaps)
        lx = (len0 < len1)? len0 : len1;
    else
        lx = (lx < ly)? lx : ly;
    pct = 100.*(double)nm/(double)lx;
    fprintf(fx, "\n");
    fprintf(fx, "<%d match%s in an overlap of %d: %.2f percent similarity\n",
        nm, (nm == 1)? "" : "es", lx, pct);

```

**getmat**

## Table 1 (cont')

```

fprintf(fx, "<gaps in first sequence: %d", gapx);
if (gapx) {
    (void) sprintf(outx, " (%d %s%s)",
        ngapx, (dna)? "base":"residue", (ngapx == 1)? "" : "s");
    fprintf(fx, "%s", outx);

    fprintf(fx, ", gaps in second sequence: %d", gapy);
    if (gapy) {
        (void) sprintf(outx, " (%d %s%s)",
            ngapy, (dna)? "base":"residue", (ngapy == 1)? "" : "s");
        fprintf(fx, "%s", outx);
    }
    if (dna)
        fprintf(fx,
            "\n<score: %d (match = %d, mismatch = %d, gap penalty = %d + %d per base)\n",
            smax, DMAT, DMIS, DINS0, DINS1);
    else
        fprintf(fx,
            "\n<score: %d (Dayhoff PAM 250 matrix, gap penalty = %d + %d per residue)\n",
            smax, PINS0, PINS1);
    if (endgaps)
        fprintf(fx,
            "<endgaps penalized. left endgap: %d %s%s, right endgap: %d %s%s\n",
            firstgap, (dna)? "base" : "residue", (firstgap == 1)? "" : "s",
            lastgap, (dna)? "base" : "residue", (lastgap == 1)? "" : "s");
    else
        fprintf(fx, "<endgaps not penalized\n");
}

static      nm;           /* matches in core -- for checking */
static      lmax;         /* lengths of stripped file names */
static      ij[2];        /* jmp index for a path */
static      nc[2];        /* number at start of current line */
static      ni[2];        /* current elem number -- for gapping */
static      siz[2];
static char *ps[2];       /* ptr to current element */
static char *po[2];       /* ptr to next output char slot */
static char out[2][P_LINE]; /* output line */
static char star[P_LINE]; /* set by stars() */

/*
 * print alignment of described in struct path pp [ ]
 */
static
pr_align()
{
    int      nn;           /* char count */
    int      more;
    register i;

    for (i = 0, lmax = 0; i < 2; i++) {
        nn = stripname(name[i]);
        if (nn > lmax)
            lmax = nn;

        nc[i] = 1;
        ni[i] = 1;
        siz[i] = ij[i] = 0;
    }
}

```

...getmat

pr\_align

## Table 1 (cont')

...pr\_align

```
    ps[i] = seqx[i];
    po[i] = out[i];
    }
    for (nn = nm = 0, more = 1; more; ) {
        for (i = more = 0; i < 2; i++) {
            /*
             * do we have more of this sequence?
             */
            if (!*ps[i])
                continue;

            more++;

            if (pp[i].spc) { /* leading space */
                *po[i]++ = ' ';
                pp[i].spc--;
            }
            else if (siz[i]) { /* in a gap */
                *po[i]++ = '-';
                siz[i]--;
            }
            else { /* we're putting a seq element
             */
                *po[i] = *ps[i];
                if (islower(*ps[i]))
                    *ps[i] = toupper(*ps[i]);
                po[i]++;
                ps[i]++;

                /*
                 * are we at next gap for this seq?
                 */
                if (ni[i] == pp[i].x[ij[i]]) {
                    /*
                     * we need to merge all gaps
                     * at this location
                     */
                    siz[i] = pp[i].n[ij[i]]++;
                    while (ni[i] == pp[i].x[ij[i]])
                        siz[i] += pp[i].n[ij[i]]++;
                }
                ni[i]++;
            }
        }
        if (++nn == olen || !more && nn) {
            dumpblock();
            for (i = 0; i < 2; i++)
                po[i] = out[i];
            nn = 0;
        }
    }
}

/*
 * dump a block of lines, including numbers, stars: pr_align()
 */
static
```

## Table 1 (cont')

dumpblock()

dumpblock

```
{
    register i;

    for (i = 0; i < 2; i++)
        *po[i]-- = '\0';
    (void) putc('\n', fx);
    for (i = 0; i < 2; i++) {
        if (*out[i] && (*out[i] != ' ' || *(po[i]) != ' ')) {
            if (i == 0)
                nums(i);
            if (i == 0 && *out[1])
                stars();
            putline(i);
            if (i == 0 && *out[1])
                fprintf(fx, star);
            if (i == 1)
                nums(i);
        }
    }
}

/*
 * put out a number line: dumpblock()
 */
static
nums(ix)
    int ix; /* index in out[] holding seq line */
{
    char nline[P_LINE];
    register i, j;
    register char *pn, *px, *py;

    for (pn = nline, i = 0; i < lmax+P_SPC; i++, pn++)
        *pn = ' ';
    for (i = nc[ix], py = out[ix]; *py; py++, pn++) {
        if (*py == ' ' || *py == '-')
            *pn = ' ';
        else {
            if (i%10 == 0 || (i == 1 && nc[ix] != 1)) {
                j = (i < 0)? -i : i;
                for (px = pn; j /= 10, px--)
                    *px = j%10 + '0';
                if (i < 0)
                    *px = '-';
            }
            else
                *pn = ' ';
            i++;
        }
    }
    *pn = '\0';
    nc[ix] = i;
    for (pn = nline; *pn; pn++)
        (void) putc(*pn, fx);
    (void) putc('\n', fx);
}

/*
```

nums

### Table 1 (cont')

**\* put out a line (name, [num], seq, [num]): dumpblock()**

**\*/**

**static**

**putline(ix)**

**putline**

```

    int      ix;
    int      i;
    register char *px;

    for (px = namex[ix], i = 0; *px && *px != ':'; px++, i++)
        (void) putc(*px, fx);
    for (; i < lmax+P_SPC; i++)
        (void) putc(' ', fx);

    /* these count from 1:
     * ni[] is current element (from 1)
     * nc[] is number at start of current line
     */
    for (px = out[ix]; *px; px++)
        (void) putc(*px&0x7F, fx);
    (void) putc('\n', fx);

```

**}**

**/\***

**\* put a line of stars (seqs always in out[0], out[1]): dumpblock()**

**\*/**

**static**

**stars()**

**stars**

**{**

```

    int      i;
    register char *p0, *p1, cx, *px;

    if (!*out[0] || (*out[0] == ' ' && *(po[0]) == ' ') ||
        !*out[1] || (*out[1] == ' ' && *(po[1]) == ' '))
        return;
    px = star;
    for (i = lmax+P_SPC; i; i--)
        *px++ = ' ';

    for (p0 = out[0], p1 = out[1]; *p0 && *p1; p0++, p1++) {
        if (isalpha(*p0) && isalpha(*p1)) {
            if (xbm[*p0-'A']&xbm[*p1-'A']) {
                cx = '*';
                nm++;
            }
            else if (!dna && _day[*p0-'A'][*p1-'A'] > 0)
                cx = ':';
            else
                cx = ' ';
        }
        else
            cx = ' ';
        *px++ = cx;
    }
    *px++ = '\n';
    *px = '\0';

```

**}**

### Table 1 (cont')

```
/*
 * strip path or prefix from pn, return len: pr_align()
 */
```

**static**

**stripname(pn)**

**stripname**

```
    char    *pn;    /* file name (may be path) */
{
    register char    *px, *py;

    py = 0;
    for (px = pn; *px; px++)
        if (*px == '/')
            py = px + 1;
    if (py)
        (void) strcpy(pn, py);
    return(strlen(pn));
}
```

```
/*
 * cleanup() -- cleanup any tmp file
 * getseq() -- read in seq, set dna, len, maxlen
 * g_calloc() -- calloc() with error checkin
 * readjumps() -- get the good jumps, from tmp file if necessary
 * writejumps() -- write a filled array of jumps to a tmp file: nw()
 */
```

**#include "nw.h"**

**#include <sys/file.h>**

```
char    *jname = "/tmp/homgXXXXXX";    /* tmp file for jumps */
FILE    *fj;
```

```
int    cleanup();    /* cleanup tmp file */
long    lseek();
```

```
/*
 * remove any tmp file if we blow
 */
```

**cleanup(i)**

**cleanup**

```
    int    i;
{
    if (fj)
        (void) unlink(jname);
    exit(i);
}
```

```
/*
 * read, return ptr to seq, set dna, len, maxlen
 * skip lines starting with ';', '<', or '>'
 * seq in upper or lower case
 */
char    *
```



**Table 1 (cont')**

getseq(file, len)

getseq

```

    char    *file;    /* file name */
    int     *len;     /* seq len */
{
    char    line[1024], *pseq;
    register char *px, *py;
    int     natgc, tlen;
    FILE    *fp;

    if ((fp = fopen(file, "r")) == 0) {
        fprintf(stderr, "%s: can't read %s\n", prog, file);
        exit(1);
    }
    tlen = natgc = 0;
    while (fgets(line, 1024, fp)) {
        if (*line == ';' || *line == '<' || *line == '>')
            continue;
        for (px = line; *px != '\n'; px++)
            if (isupper(*px) || islower(*px))
                tlen++;
    }
    if ((pseq = malloc((unsigned)(tlen+6))) == 0) {
        fprintf(stderr, "%s: malloc() failed to get %d bytes for %s\n", prog, tlen+6, file);
        exit(1);
    }
    pseq[0] = pseq[1] = pseq[2] = pseq[3] = '\0';
    py = pseq + 4;
    *len = tlen;
    rewind(fp);

    while (fgets(line, 1024, fp)) {
        if (*line == ';' || *line == '<' || *line == '>')
            continue;
        for (px = line; *px != '\n'; px++) {
            if (isupper(*px))
                *py++ = *px;
            else if (islower(*px))
                *py++ = toupper(*px);
            if (index("ATGCU", *(py-1)))
                natgc++;
        }
        *py++ = '\0';
        *py = '\0';
    }
    (void) fclose(fp);
    dna = natgc > (tlen/3);
    return(pseq+4);
}

char    *
```

## Table 1 (cont')

<pre> g_alloc(msg, nx, sz)     char    *msg;          /* program, calling routine */     int      nx, sz;        /* number and size of elements */ {     char      *px, *calloc();      if ((px = calloc((unsigned)nx, (unsigned)sz)) == 0) {         if (*msg) {             fprintf(stderr, "%s: g_alloc() failed %s (n=%d, sz=%d)\n", prog, msg, nx, sz);             exit(1);         }     }     return(px); } </pre>	<p><b>g_alloc</b></p>
<pre> /*  * get final jmps from dx[] or tmp file, set pp[], reset dmax: main()  */ readjmps() {     int      fd = -1;     int      siz, i0, i1;     register i, j, xx;      if (fj) {         (void) fclose(fj);         if ((fd = open(jname, O_RDONLY, 0)) &lt; 0) {             fprintf(stderr, "%s: can't open() %s\n", prog, jname);             cleanup(1);         }     }     for (i = i0 = i1 = 0, dmax0 = dmax, xx = len0; ; i++) {         while (1) {             for (j = dx[dmax].ijmp; j &gt;= 0 &amp;&amp; dx[dmax].jp.x[j] &gt;= xx; j--)                 ;             if (j &lt; 0 &amp;&amp; dx[dmax].offset &amp;&amp; fj) {                 (void) lseek(fd, dx[dmax].offset, 0);                 (void) read(fd, (char *)&amp;dx[dmax].jp, sizeof(struct jmp));                 (void) read(fd, (char *)&amp;dx[dmax].offset, sizeof(dx[dmax].offset));                 dx[dmax].ijmp = MAXJMP-1;             }             else                 break;         }         if (i &gt;= JMPS) {             fprintf(stderr, "%s: too many gaps in alignment\n", prog);             cleanup(1);         }     } } </pre>	<p><b>readjmps</b></p>

## Table 1 (cont')

...readjumps

```
    }
    if (j >= 0) {
        siz = dx[dmax].jp.n[j];
        xx = dx[dmax].jp.x[j];
        dmax += siz;
        if (siz < 0) { /* gap in second seq */
            pp[1].n[i1] = -siz;
            xx += siz;
            /* id = xx - yy + len1 - 1
            */
            pp[1].x[i1] = xx - dmax + len1 - 1;
            gapy++;
            ngapy -= siz;
            /* ignore MAXGAP when doing endgaps */
            siz = (-siz < MAXGAP || endgaps)? -siz : MAXGAP;
            i1++;
        }
        else if (siz > 0) { /* gap in first seq */
            pp[0].n[i0] = siz;
            pp[0].x[i0] = xx;
            gapx++;
            ngapx += siz;
            /* ignore MAXGAP when doing endgaps */
            siz = (siz < MAXGAP || endgaps)? siz : MAXGAP;
            i0++;
        }
    }
    else
        break;
}

/* reverse the order of jumps
*/
for (j = 0, i0--; j < i0; j++, i0--) {
    i = pp[0].n[j]; pp[0].n[j] = pp[0].n[i0]; pp[0].n[i0] = i;
    i = pp[0].x[j]; pp[0].x[j] = pp[0].x[i0]; pp[0].x[i0] = i;
}
for (j = 0, i1--; j < i1; j++, i1--) {
    i = pp[1].n[j]; pp[1].n[j] = pp[1].n[i1]; pp[1].n[i1] = i;
    i = pp[1].x[j]; pp[1].x[j] = pp[1].x[i1]; pp[1].x[i1] = i;
}
if (fd >= 0)
    (void) close(fd);
if (fj) {
    (void) unlink(jname);
}
```

### Table 1 (cont')

```
        fj = 0;
        offset = 0;
    }

/*
 * write a filled jmp struct offset of the prev one (if any): nw()
 */
writejmps(ix)
    int    ix;
{
    char    *mktemp();

    if (!fj) {
        if (mktemp(jname) < 0) {
            fprintf(stderr, "%s: can't mktemp() %s\n", prog, jname);
            cleanup(1);
        }
        if ((fj = fopen(jname, "w")) == 0) {
            fprintf(stderr, "%s: can't write %s\n", prog, jname);
            exit(1);
        }
    }
    (void) fwrite((char *)&dx[ix].jp, sizeof(struct jmp), 1, fj);
    (void) fwrite((char *)&dx[ix].offset, sizeof(dx[ix].offset), 1, fj);
}
```

**writejmps**

**Table 2**

PRO	XXXXXXXXXXXXXXXXX	(Length = 15 amino acids)
Comparison Protein	XXXXXXXXXXXXXX	(Length = 12 amino acids)

% amino acid sequence identity =

(the number of identically matching amino acid residues between the two polypeptide sequences as determined by ALIGN-2) divided by (the total number of amino acid residues of the PRO polypeptide)  
=

5 divided by 15 = 33.3%

**Table 3**

PRO	XXXXXXXXXX	(Length = 10 amino acids)
Comparison Protein	XXXXXXXXXXXXZZYZ	(Length = 15 amino acids)

% amino acid sequence identity =

(the number of identically matching amino acid residues between the two polypeptide sequences as determined by ALIGN-2) divided by (the total number of amino acid residues of the PRO polypeptide)  
=

5 divided by 10 = 50%

**Table 4**

PRO-DNA	NNNNNNNNNNNNNN	(Length = 14 nucleotides)
Comparison DNA	NNNNNNLLLLLLLLLL	(Length = 16 nucleotides)

% nucleic acid sequence identity =

(the number of identically matching nucleotides between the two nucleic acid sequences as determined by ALIGN-2) divided by (the total number of nucleotides of the PRO-DNA nucleic acid sequence) =

6 divided by 14 = 42.9%

**Table 5**

PRO-DNA	NNNNNNNNNNNN	(Length = 12 nucleotides)
Comparison DNA	NNNNLLLLV	(Length = 9 nucleotides)

% nucleic acid sequence identity =

(the number of identically matching nucleotides between the two nucleic acid sequences as determined by ALIGN-2) divided by (the total number of nucleotides of the PRO-DNA nucleic acid sequence) =

4 divided by 12 = 33.3%

## II. Compositions and Methods of the Invention

### A. Full-Length PRO Polypeptides

[0244] The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO polypeptides. In particular, cDNAs encoding various PRO polypeptides have been identified and isolated, as disclosed in further detail in the Examples below. It is noted that proteins produced in separate expression rounds may be given different PRO numbers but the UNQ number is unique for any given DNA and the encoded protein, and will not be changed. However, for sake of simplicity, in the present specification the protein encoded by the full length native nucleic acid molecules disclosed herein as well as all further native homologues and variants included in the foregoing definition of PRO, will be referred to as “PRO/number”, regardless of their origin or mode of preparation.

[0245] As disclosed in the Examples below, various cDNA clones have been deposited with the ATCC. The actual nucleotide sequences of those clones can readily be determined by the skilled artisan by sequencing of the deposited clone using routine methods in the art. The predicted amino acid sequence can be determined from the nucleotide sequence using routine skill. For the PRO polypeptides and encoding nucleic acids described herein, Applicants have identified what is believed to be the reading frame best identifiable with the sequence information available at the time.

### B. PRO Polypeptide Variants

[0246] In addition to the full-length native sequence PRO polypeptides described herein, it is contemplated that PRO variants can be prepared. PRO variants can be prepared by introducing appropriate nucleotide changes into the PRO DNA, and/or by synthesis of the desired PRO polypeptide. Those skilled in the art will appreciate that amino acid changes may alter post-translational processes of the PRO, such as changing the number or position of glycosylation sites or altering the membrane anchoring characteristics.

[0247] Variations in the native full-length sequence PRO or in various domains of the PRO described herein, can be made, for example, using any of the techniques and guidelines for conservative and non-conservative mutations set forth, for instance, in U.S.

Patent No. 5,364,934. Variations may be a substitution, deletion or insertion of one or more codons encoding the PRO that results in a change in the amino acid sequence of the PRO as compared with the native sequence PRO. Optionally the variation is by substitution of at least one amino acid with any other amino acid in one or more of the domains of the PRO. Guidance in determining which amino acid residue may be inserted, substituted or deleted without adversely affecting the desired activity may be found by comparing the sequence of the PRO with that of homologous known protein molecules and minimizing the number of amino acid sequence changes made in regions of high homology. Amino acid substitutions can be the result of replacing one amino acid with another amino acid having similar structural and/or chemical properties, such as the replacement of a leucine with a serine, i.e., conservative amino acid replacements. Insertions or deletions may optionally be in the range of about 1 to 5 amino acids. The variation allowed may be determined by systematically making insertions, deletions or substitutions of amino acids in the sequence and testing the resulting variants for activity exhibited by the full-length or mature native sequence.

**[0248]** PRO polypeptide fragments are provided herein. Such fragments may be truncated at the N-terminus or C-terminus, or may lack internal residues, for example, when compared with a full length native protein. Certain fragments lack amino acid residues that are not essential for a desired biological activity of the PRO polypeptide.

**[0249]** PRO fragments may be prepared by any of a number of conventional techniques. Desired peptide fragments may be chemically synthesized. An alternative approach involves generating PRO fragments by enzymatic digestion, e.g., by treating the protein with an enzyme known to cleave proteins at sites defined by particular amino acid residues, or by digesting the DNA with suitable restriction enzymes and isolating the desired fragment. Yet another suitable technique involves isolating and amplifying a DNA fragment encoding a desired polypeptide fragment, by polymerase chain reaction (PCR). Oligonucleotides that define the desired termini of the DNA fragment are employed at the 5' and 3' primers in the PCR. Preferably, PRO polypeptide fragments share at least one biological and/or immunological activity with the native PRO polypeptide disclosed herein.

**[0250]** In particular embodiments, conservative substitutions of interest are shown in Table 6 under the heading of preferred substitutions. If such substitutions result in a change



in biological activity, then more substantial changes, denominated exemplary substitutions in Table 6, or as further described below in reference to amino acid classes, are introduced and the products screened.

Table 6

Original Residue	Exemplary Substitutions	Preferred Substitutions
Ala (A)	val; leu; ile	val
Arg (R)	lys; gln; asn	lys
Asn (N)	gln; his; lys; arg	gln
Asp (D)	glu	glu
Cys (C)	ser	ser
Gln (Q)	asn	asn
Glu (E)	asp	asp
Gly (G)	pro; ala	ala
His (H)	asn; gln; lys; arg	arg
Ile (I)	leu; val; met; ala; phe; norleucine	leu
Leu (L)	norleucine; ile; val; met; ala; phe	ile
Lys (K)	arg; gln; asn	arg
Met (M)	leu; phe; ile	leu
Phe (F)	leu; val; ile; ala; tyr	leu
Pro (P)	ala	ala
Ser (S)	thr	thr
Thr (T)	ser	ser
Trp (W)	tyr; phe	tyr
Tyr (Y)	trp; phe; thr; ser	phe
Val (V)	ile; leu; met; phe; ala; norleucine	leu

**[0251]** Substantial modifications in function or immunological identity of the PRO polypeptide are accomplished by selecting substitutions that differ significantly in their effect on maintaining (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a sheet or helical conformation, (b) the charge or hydrophobicity of the molecule at the target site, or (c) the bulk of the side chain. Naturally occurring residues are divided into groups based on common side-chain properties:

- (1) hydrophobic: norleucine, met, ala, val, leu, ile;
- (2) neutral hydrophilic: cys, ser, thr;
- (3) acidic: asp, glu;
- (4) basic: asn, gln, his, lys, arg;

(5) residues that influence chain orientation: gly, pro; and

(6) aromatic: trp, tyr, phe.

[0252] Non-conservative substitutions will entail exchanging a member of one of these classes for another class. Such substituted residues also may be introduced into the conservative substitution sites or, more preferably, into the remaining (non-conserved) sites.

[0253] The variations can be made using methods known in the art such as oligonucleotide-mediated (site-directed) mutagenesis, alanine scanning, and PCR mutagenesis. Site-directed mutagenesis [Carter et al., Nucl. Acids Res., 13:4331 (1986); Zoller et al., Nucl. Acids Res., 10:6487 (1987)], cassette mutagenesis [Wells et al., Gene, 34:315 (1985)], restriction selection mutagenesis [Wells et al., Philos. Trans. R. Soc. London SerA, 317:415 (1986)] or other known techniques can be performed on the cloned DNA to produce the PRO variant DNA.

[0254] Scanning amino acid analysis can also be employed to identify one or more amino acids along a contiguous sequence. Among the preferred scanning amino acids are relatively small, neutral amino acids. Such amino acids include alanine, glycine, serine, and cysteine. Alanine is typically a preferred scanning amino acid among this group because it eliminates the side-chain beyond the beta-carbon and is less likely to alter the main-chain conformation of the variant [Cunningham and Wells, Science, 244: 1081-1085 (1989)]. Alanine is also typically preferred because it is the most common amino acid. Further, it is frequently found in both buried and exposed positions [Creighton, The Proteins, (W.H. Freeman & Co., N.Y.); Chothia, J. Mol. Biol., 150:1 (1976)]. If alanine substitution does not yield adequate amounts of variant, an isoteric amino acid can be used.

### C. Modifications of PRO

[0255] Covalent modifications of PRO are included within the scope of this invention. One type of covalent modification includes reacting targeted amino acid residues of a PRO polypeptide with an organic derivatizing agent that is capable of reacting with selected side chains or the N- or C- terminal residues of the PRO. Derivatization with bifunctional agents is useful, for instance, for crosslinking PRO to a water-insoluble support matrix or surface for use in the method for purifying anti-PRO antibodies, and vice-versa. Commonly

used crosslinking agents include, e.g., 1,1-bis(diazoacetyl)-2-phenylethane, glutaraldehyde, N-hydroxysuccinimide esters, for example, esters with 4-azidosalicylic acid, homobifunctional imidoesters, including disuccinimidyl esters such as 3,3'-dithiobis(succinimidylpropionate), bifunctional maleimides such as bis-N-maleimido-1,8-octane and agents such as methyl-3-[(p-azidophenyl)dithio]propioimide.

[0256] Other modifications include deamidation of glutamyl and asparaginyl residues to the corresponding glutamyl and aspartyl residues, respectively, hydroxylation of proline and lysine, phosphorylation of hydroxyl groups of seryl or threonyl residues, methylation of the  $\alpha$ -amino groups of lysine, arginine, and histidine side chains [T.E. Creighton, Proteins: Structure and Molecular Properties, W.H. Freeman & Co., San Francisco, pp. 79-86 (1983)], acetylation of the N-terminal amine, and amidation of any C-terminal carboxyl group.

[0257] Another type of covalent modification of the PRO polypeptide included within the scope of this invention comprises altering the native glycosylation pattern of the polypeptide. "Altering the native glycosylation pattern" is intended for purposes herein to mean deleting one or more carbohydrate moieties found in native sequence PRO (either by removing the underlying glycosylation site or by deleting the glycosylation by chemical and/or enzymatic means), and/or adding one or more glycosylation sites that are not present in the native sequence PRO. In addition, the phrase includes qualitative changes in the glycosylation of the native proteins, involving a change in the nature and proportions of the various carbohydrate moieties present.

[0258] Addition of glycosylation sites to the PRO polypeptide may be accomplished by altering the amino acid sequence. The alteration may be made, for example, by the addition of, or substitution by, one or more serine or threonine residues to the native sequence PRO (for O-linked glycosylation sites). The PRO amino acid sequence may optionally be altered through changes at the DNA level, particularly by mutating the DNA encoding the PRO polypeptide at preselected bases such that codons are generated that will translate into the desired amino acids.

[0259] Another means of increasing the number of carbohydrate moieties on the PRO polypeptide is by chemical or enzymatic coupling of glycosides to the polypeptide. Such

methods are described in the art, e.g., in WO 87/05330 published 11 September 1987, and in Aplin and Wriston, CRC Crit. Rev. Biochem., pp. 259-306 (1981).

[0260] Removal of carbohydrate moieties present on the PRO polypeptide may be accomplished chemically or enzymatically or by mutational substitution of codons encoding for amino acid residues that serve as targets for glycosylation. Chemical deglycosylation techniques are known in the art and described, for instance, by Hakimuddin, et al., Arch. Biochem. Biophys., 259:52 (1987) and by Edge et al., Anal. Biochem., 118:131 (1981). Enzymatic cleavage of carbohydrate moieties on polypeptides can be achieved by the use of a variety of endo- and exo-glycosidases as described by Thotakura et al., Meth. Enzymol., 138:350 (1987).

[0261] Another type of covalent modification of PRO comprises linking the PRO polypeptide to one of a variety of nonproteinaceous polymers, e.g., polyethylene glycol (PEG), polypropylene glycol, or polyoxyalkylenes, in the manner set forth in U.S. Patent Nos. 4,640,835; 4,496,689; 4,301,144; 4,670,417; 4,791,192 or 4,179,337.

[0262] The PRO of the present invention may also be modified in a way to form a chimeric molecule comprising PRO fused to another, heterologous polypeptide or amino acid sequence.

[0263] In one embodiment, such a chimeric molecule comprises a fusion of the PRO with a tag polypeptide which provides an epitope to which an anti-tag antibody can selectively bind. The epitope tag is generally placed at the amino- or carboxyl- terminus of the PRO. The presence of such epitope-tagged forms of the PRO can be detected using an antibody against the tag polypeptide. Also, provision of the epitope tag enables the PRO to be readily purified by affinity purification using an anti-tag antibody or another type of affinity matrix that binds to the epitope tag. Various tag polypeptides and their respective antibodies are well known in the art. Examples include poly-histidine (poly-his) or poly-histidine-glycine (poly-his-gly) tags; the flu HA tag polypeptide and its antibody 12CA5 [Field et al., Mol. Cell. Biol., 8:2159-2165 (1988)]; the c-myc tag and the 8F9, 3C7, 6E10, G4, B7 and 9E10 antibodies thereto [Evan et al., Molecular and Cellular Biology, 5:3610-3616 (1985)]; and the Herpes Simplex virus glycoprotein D (gD) tag and its antibody [Paborsky et al., Protein Engineering, 3(6):547-553 (1990)]. Other tag polypeptides include the Flag-peptide [Hopp et

al., BioTechnology, 6:1204-1210 (1988)]; the KT3 epitope peptide [Martin et al., Science, 255:192-194 (1992)]; an  $\alpha$ -tubulin epitope peptide [Skinner et al., J. Biol. Chem., 266:15163-15166 (1991)]; and the T7 gene 10 protein peptide tag [Lutz-Freyermuth et al., Proc. Natl. Acad. Sci. USA, 87:6393-6397 (1990)].

[0264] In an alternative embodiment, the chimeric molecule may comprise a fusion of the PRO with an immunoglobulin or a particular region of an immunoglobulin. For a bivalent form of the chimeric molecule (also referred to as an “immunoadhesin”), such a fusion could be to the Fc region of an IgG molecule. The Ig fusions preferably include the substitution of a soluble (transmembrane domain deleted or inactivated) form of a PRO polypeptide in place of at least one variable region within an Ig molecule. In a particularly preferred embodiment, the immunoglobulin fusion includes the hinge, CH2 and CH3, or the hinge, CH1, CH2 and CH3 regions of an IgG1 molecule. For the production of immunoglobulin fusions see also US Patent No. 5,428,130 issued June 27, 1995.

#### D. Preparation of PRO

[0265] The description below relates primarily to production of PRO by culturing cells transformed or transfected with a vector containing PRO nucleic acid. It is, of course, contemplated that alternative methods, which are well known in the art, may be employed to prepare PRO. For instance, the PRO sequence, or portions thereof, may be produced by direct peptide synthesis using solid-phase techniques [see, e.g., Stewart et al., Solid-Phase Peptide Synthesis, W.H. Freeman Co., San Francisco, CA (1969); Merrifield, J. Am. Chem. Soc., 85:2149-2154 (1963)]. *In vitro* protein synthesis may be performed using manual techniques or by automation. Automated synthesis may be accomplished, for instance, using an Applied Biosystems Peptide Synthesizer (Foster City, CA) using manufacturer’s instructions. Various portions of the PRO may be chemically synthesized separately and combined using chemical or enzymatic methods to produce the full-length PRO.

##### 1. Isolation of DNA Encoding PRO

[0266] DNA encoding PRO may be obtained from a cDNA library prepared from tissue believed to possess the PRO mRNA and to express it at a detectable level. Accordingly,

human PRO DNA can be conveniently obtained from a cDNA library prepared from human tissue, such as described in the Examples. The PRO-encoding gene may also be obtained from a genomic library or by known synthetic procedures (e.g., automated nucleic acid synthesis).

[0267] Libraries can be screened with probes (such as antibodies to the PRO or oligonucleotides of at least about 20-80 bases) designed to identify the gene of interest or the protein encoded by it. Screening the cDNA or genomic library with the selected probe may be conducted using standard procedures, such as described in Sambrook et al., Molecular Cloning: A Laboratory Manual (New York: Cold Spring Harbor Laboratory Press, 1989). An alternative means to isolate the gene encoding PRO is to use PCR methodology [Sambrook et al., supra; Dieffenbach et al., PCR Primer: A Laboratory Manual (Cold Spring Harbor Laboratory Press, 1995)].

[0268] The Examples below describe techniques for screening a cDNA library. The oligonucleotide sequences selected as probes should be of sufficient length and sufficiently unambiguous that false positives are minimized. The oligonucleotide is preferably labeled such that it can be detected upon hybridization to DNA in the library being screened. Methods of labeling are well known in the art, and include the use of radiolabels like <sup>32</sup>P-labeled ATP, biotinylation or enzyme labeling. Hybridization conditions, including moderate stringency and high stringency, are provided in Sambrook et al., supra.

[0269] Sequences identified in such library screening methods can be compared and aligned to other known sequences deposited and available in public databases such as GenBank or other private sequence databases. Sequence identity (at either the amino acid or nucleotide level) within defined regions of the molecule or across the full-length sequence can be determined using methods known in the art and as described herein.

[0270] Nucleic acid having protein coding sequence may be obtained by screening selected cDNA or genomic libraries using the deduced amino acid sequence disclosed herein for the first time, and, if necessary, using conventional primer extension procedures as described in Sambrook et al., supra, to detect precursors and processing intermediates of mRNA that may not have been reverse-transcribed into cDNA.

## 2. Selection and Transformation of Host Cells

[0271] Host cells are transfected or transformed with expression or cloning vectors described herein for PRO production and cultured in conventional nutrient media modified as appropriate for inducing promoters, selecting transformants, or amplifying the genes encoding the desired sequences. The culture conditions, such as media, temperature, pH and the like, can be selected by the skilled artisan without undue experimentation. In general, principles, protocols, and practical techniques for maximizing the productivity of cell cultures can be found in Mammalian Cell Biotechnology: a Practical Approach, M. Butler, ed. (IRL Press, 1991) and Sambrook et al., supra.

[0272] Methods of eukaryotic cell transfection and prokaryotic cell transformation are known to the ordinarily skilled artisan, for example,  $\text{CaCl}_2$ ,  $\text{CaPO}_4$ , liposome-mediated and electroporation. Depending on the host cell used, transformation is performed using standard techniques appropriate to such cells. The calcium treatment employing calcium chloride, as described in Sambrook et al., supra, or electroporation is generally used for prokaryotes. Infection with *Agrobacterium tumefaciens* is used for transformation of certain plant cells, as described by Shaw et al., Gene, 23:315 (1983) and WO 89/05859 published 29 June 1989. For mammalian cells without such cell walls, the calcium phosphate precipitation method of Graham and van der Eb, Virology, 52:456-457 (1978) can be employed. General aspects of mammalian cell host system transfections have been described in U.S. Patent No. 4,399,216. Transformations into yeast are typically carried out according to the method of Van Solingen et al., J. Bact., 130:946 (1977) and Hsiao et al., Proc. Natl. Acad. Sci. (USA), 76:3829 (1979). However, other methods for introducing DNA into cells, such as by nuclear microinjection, electroporation, bacterial protoplast fusion with intact cells, or polycations, e.g., polybrene, polyornithine, may also be used. For various techniques for transforming mammalian cells, see Keown et al., Methods in Enzymology, 185:527-537 (1990) and Mansour et al., Nature, 336:348-352 (1988).

[0273] Suitable host cells for cloning or expressing the DNA in the vectors herein include prokaryote, yeast, or higher eukaryote cells. Suitable prokaryotes include but are not limited to eubacteria, such as Gram-negative or Gram-positive organisms, for example, Enterobacteriaceae such as *E. coli*. Various *E. coli* strains are publicly available, such as *E. coli*

K12 strain MM294 (ATCC 31,446); *E. coli* X1776 (ATCC 31,537); *E. coli* strain W3110 (ATCC 27,325) and K5 772 (ATCC 53,635). Other suitable prokaryotic host cells include Enterobacteriaceae such as *Escherichia*, e.g., *E. coli*, *Enterobacter*, *Erwinia*, *Klebsiella*, *Proteus*, *Salmonella*, e.g., *Salmonella typhimurium*, *Serratia*, e.g., *Serratia marcescans*, and *Shigella*, as well as *Bacilli* such as *B. subtilis* and *B. licheniformis* (e.g., *B. licheniformis* 41P disclosed in DD 266,710 published 12 April 1989), *Pseudomonas* such as *P. aeruginosa*, and *Streptomyces*. These examples are illustrative rather than limiting. Strain W3110 is one particularly preferred host or parent host because it is a common host strain for recombinant DNA product fermentations. Preferably, the host cell secretes minimal amounts of proteolytic enzymes. For example, strain W3110 may be modified to effect a genetic mutation in the genes encoding proteins endogenous to the host, with examples of such hosts including *E. coli* W3110 strain 1A2, which has the complete genotype *tonA*; *E. coli* W3110 strain 9E4, which has the complete genotype *tonA ptr3*; *E. coli* W3110 strain 27C7 (ATCC 55,244), which has the complete genotype *tonA ptr3 phoA E15 (argF-lac)169 degP ompT kan<sup>r</sup>*; *E. coli* W3110 strain 37D6, which has the complete genotype *tonA ptr3 phoA E15 (argF-lac)169 degP ompT rbs7 ilvG kan<sup>r</sup>*; *E. coli* W3110 strain 40B4, which is strain 37D6 with a non-kanamycin resistant *degP* deletion mutation; and an *E. coli* strain having mutant periplasmic protease disclosed in U.S. Patent No. 4,946,783 issued 7 August 1990. Alternatively, *in vitro* methods of cloning, e.g., PCR or other nucleic acid polymerase reactions, are suitable.

[0274] In addition to prokaryotes, eukaryotic microbes such as filamentous fungi or yeast are suitable cloning or expression hosts for PRO-encoding vectors. *Saccharomyces cerevisiae* is a commonly used lower eukaryotic host microorganism. Others include *Schizosaccharomyces pombe* (Beach and Nurse, Nature, 290: 140 [1981]; EP 139,383 published 2 May 1985); *Kluyveromyces* hosts (U.S. Patent No. 4,943,529; Fleer et al., Bio/Technology, 9:968-975 (1991)) such as, e.g., *K. lactis* (MW98-8C, CBS683, CBS4574; Louvencourt et al., J. Bacteriol., 154(2):737-742 [1983]), *K. fragilis* (ATCC 12,424), *K. bulgaricus* (ATCC 16,045), *K. wickerhamii* (ATCC 24,178), *K. waltii* (ATCC 56,500), *K. drosophilum* (ATCC 36,906; Van den Berg et al., Bio/Technology, 8:135 (1990)), *K. thermotolerans*, and *K. marxianus*; *yarrowia* (EP 402,226); *Pichia pastoris* (EP 183,070; Sreekrishna et al., J. Basic Microbiol., 28:265-278 [1988]); *Candida*; *Trichoderma reesia* (EP



244,234); *Neurospora crassa* (Case et al., Proc. Natl. Acad. Sci. USA, 76:5259-5263 [1979]); *Schwanniomyces* such as *Schwanniomyces occidentalis* (EP 394,538 published 31 October 1990); and filamentous fungi such as, e.g., *Neurospora*, *Penicillium*, *Tolypocladium* (WO 91/00357 published 10 January 1991), and *Aspergillus* hosts such as *A. nidulans* (Ballance et al., Biochem. Biophys. Res. Commun., 112:284-289 [1983]; Tilburn et al., Gene, 26:205-221 [1983]; Yelton et al., Proc. Natl. Acad. Sci. USA, 81: 1470-1474 [1984]) and *A. niger* (Kelly and Hynes, EMBO J., 4:475-479 [1985]). Methylotropic yeasts are suitable herein and include, but are not limited to, yeast capable of growth on methanol selected from the genera consisting of *Hansenula*, *Candida*, *Kloeckera*, *Pichia*, *Saccharomyces*, *Torulopsis*, and *Rhodotorula*. A list of specific species that are exemplary of this class of yeasts may be found in C. Anthony, The Biochemistry of Methylotrophs, 269 (1982).

[0275] Suitable host cells for the expression of glycosylated PRO are derived from multicellular organisms. Examples of invertebrate cells include insect cells such as *Drosophila* S2 and *Spodoptera Sf9*, as well as plant cells. Examples of useful mammalian host cell lines include Chinese hamster ovary (CHO) and COS cells. More specific examples include monkey kidney CV1 line transformed by SV40 (COS-7, ATCC CRL 1651); human embryonic kidney line (293 or 293 cells subcloned for growth in suspension culture, Graham et al., J. Gen Virol., 36:59 (1977)); Chinese hamster ovary cells/-DHFR (CHO, Urlaub and Chasin, Proc. Natl. Acad. Sci. USA, 77:4216 (1980)); mouse sertoli cells (TM4, Mather, Biol. Reprod., 23:243-251 (1980)); human lung cells (W138, ATCC CCL 75); human liver cells (Hep G2, HB 8065); and mouse mammary tumor (MMT 060562, ATCC CCL51). The selection of the appropriate host cell is deemed to be within the skill in the art.

### 3. Selection and Use of a Replicable Vector

[0276] The nucleic acid (e.g., cDNA or genomic DNA) encoding PRO may be inserted into a replicable vector for cloning (amplification of the DNA) or for expression. Various vectors are publicly available. The vector may, for example, be in the form of a plasmid, cosmid, viral particle, or phage. The appropriate nucleic acid sequence may be inserted into the vector by a variety of procedures. In general, DNA is inserted into an appropriate restriction endonuclease site(s) using techniques known in the art. Vector

components generally include, but are not limited to, one or more of a signal sequence, an origin of replication, one or more marker genes, an enhancer element, a promoter, and a transcription termination sequence. Construction of suitable vectors containing one or more of these components employs standard ligation techniques which are known to the skilled artisan.

[0277] The PRO may be produced recombinantly not only directly, but also as a fusion polypeptide with a heterologous polypeptide, which may be a signal sequence or other polypeptide having a specific cleavage site at the N-terminus of the mature protein or polypeptide. In general, the signal sequence may be a component of the vector, or it may be a part of the PRO-encoding DNA that is inserted into the vector. The signal sequence may be a prokaryotic signal sequence selected, for example, from the group of the alkaline phosphatase, penicillinase, lpp, or heat-stable enterotoxin II leaders. For yeast secretion the signal sequence may be, e.g., the yeast invertase leader, alpha factor leader (including *Saccharomyces* and *Kluyveromyces* a-factor leaders, the latter described in U.S. Patent No. 5,010,182), or acid phosphatase leader, the *C. albicans* glucoamylase leader (EP 362,179 published 4 April 1990), or the signal described in WO 90/13646 published 15 November 1990. In mammalian cell expression, mammalian signal sequences may be used to direct secretion of the protein, such as signal sequences from secreted polypeptides of the same or related species, as well as viral secretory leaders.

[0278] Both expression and cloning vectors contain a nucleic acid sequence that enables the vector to replicate in one or more selected host cells. Such sequences are well known for a variety of bacteria, yeast, and viruses. The origin of replication from the plasmid pBR322 is suitable for most Gram-negative bacteria, the 2 $\mu$  plasmid origin is suitable for yeast, and various viral origins (SV40, polyoma, adenovirus, VSV or BPV) are useful for cloning vectors in mammalian cells.

[0279] Expression and cloning vectors will typically contain a selection gene, also termed a selectable marker. Typical selection genes encode proteins that (a) confer resistance to antibiotics or other toxins, e.g., ampicillin, neomycin, methotrexate, or tetracycline, (b) complement auxotrophic deficiencies, or (c) supply critical nutrients not available from complex media, e.g., the gene encoding D-alanine racemase for *Bacilli*.

[0280] An example of suitable selectable markers for mammalian cells are those that enable the identification of cells competent to take up the PRO-encoding nucleic acid, such as DHFR or thymidine kinase. An appropriate host cell when wild-type DHFR is employed is the CHO cell line deficient in DHFR activity, prepared and propagated as described by Urlaub et al., Proc. Natl. Acad. Sci. USA, 77:4216 (1980). A suitable selection gene for use in yeast is the *trp1* gene present in the yeast plasmid YRp7 [Stinchcomb et al., Nature, 282:39 (1979); Kingsman et al., Gene, 7:141 (1979); Tschemper et al., Gene, 10:157 (1980)]. The *trp1* gene provides a selection marker for a mutant strain of yeast lacking the ability to grow in tryptophan, for example, ATCC No. 44076 or PEP4-1 [Jones, Genetics, 85:12 (1977)].

[0281] Expression and cloning vectors usually contain a promoter operably linked to the PRO-encoding nucleic acid sequence to direct mRNA synthesis. Promoters recognized by a variety of potential host cells are well known. Promoters suitable for use with prokaryotic hosts include the  $\beta$ -lactamase and lactose promoter systems [Chang et al., Nature, 275:615 (1978); Goeddel et al., Nature, 281:544 (1979)], alkaline phosphatase, a tryptophan (*trp*) promoter system [Goeddel, Nucleic Acids Res., 8:4057 (1980); EP 36,776], and hybrid promoters such as the *tac* promoter [deBoer et al., Proc. Natl. Acad. Sci. USA, 80:21-25 (1983)]. Promoters for use in bacterial systems also will contain a Shine-Dalgarno (S.D.) sequence operably linked to the DNA encoding PRO.

[0282] Examples of suitable promoting sequences for use with yeast hosts include the promoters for 3-phosphoglycerate kinase [Hitzeman et al., J. Biol. Chem., 255:2073 (1980)] or other glycolytic enzymes [Hess et al., J. Adv. Enzyme Reg., 7:149 (1968); Holland, Biochemistry, 17:4900 (1978)], such as enolase, glyceraldehyde-3-phosphate dehydrogenase, hexokinase, pyruvate decarboxylase, phosphofructokinase, glucose-6-phosphate isomerase, 3-phosphoglycerate mutase, pyruvate kinase, triosephosphate isomerase, phosphoglucose isomerase, and glucokinase.

[0283] Other yeast promoters, which are inducible promoters having the additional advantage of transcription controlled by growth conditions, are the promoter regions for alcohol dehydrogenase 2, isocytochrome C, acid phosphatase, degradative enzymes associated with nitrogen metabolism, metallothionein, glyceraldehyde-3-phosphate dehydrogenase, and

enzymes responsible for maltose and galactose utilization. Suitable vectors and promoters for use in yeast expression are further described in EP 73,657.

**[0284]** PRO transcription from vectors in mammalian host cells is controlled, for example, by promoters obtained from the genomes of viruses such as polyoma virus, fowlpox virus (UK 2,211,504 published 5 July 1989), adenovirus (such as Adenovirus 2), bovine papilloma virus, avian sarcoma virus, cytomegalovirus, a retrovirus, hepatitis-B virus and Simian Virus 40 (SV40), from heterologous mammalian promoters, e.g., the actin promoter or an immunoglobulin promoter, and from heat-shock promoters, provided such promoters are compatible with the host cell systems.

**[0285]** Transcription of a DNA encoding the PRO by higher eukaryotes may be increased by inserting an enhancer sequence into the vector. Enhancers are cis-acting elements of DNA, usually about from 10 to 300 bp, that act on a promoter to increase its transcription. Many enhancer sequences are now known from mammalian genes (globin, elastase, albumin,  $\alpha$ -fetoprotein, and insulin). Typically, however, one will use an enhancer from a eukaryotic cell virus. Examples include the SV40 enhancer on the late side of the replication origin (bp 100-270), the cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers. The enhancer may be spliced into the vector at a position 5' or 3' to the PRO coding sequence, but is preferably located at a site 5' from the promoter.

**[0286]** Expression vectors used in eukaryotic host cells (yeast, fungi, insect, plant, animal, human, or nucleated cells from other multicellular organisms) will also contain sequences necessary for the termination of transcription and for stabilizing the mRNA. Such sequences are commonly available from the 5' and, occasionally 3', untranslated regions of eukaryotic or viral DNAs or cDNAs. These regions contain nucleotide segments transcribed as polyadenylated fragments in the untranslated portion of the mRNA encoding PRO.

**[0287]** Still other methods, vectors, and host cells suitable for adaptation to the synthesis of PRO in recombinant vertebrate cell culture are described in Gething et al., Nature, 293:620-625 (1981); Mantei et al., Nature, 281:40-46 (1979); EP 117,060; and EP 117,058.

#### 4. Detecting Gene Amplification/Expression

[0288] Gene amplification and/or expression may be measured in a sample directly, for example, by conventional Southern blotting, Northern blotting to quantitate the transcription of mRNA [Thomas, Proc. Natl. Acad. Sci. USA, 77:5201-5205 (1980)], dot blotting (DNA analysis), or *in situ* hybridization, using an appropriately labeled probe, based on the sequences provided herein. Alternatively, antibodies may be employed that can recognize specific duplexes, including DNA duplexes, RNA duplexes, and DNA-RNA hybrid duplexes or DNA-protein duplexes. The antibodies in turn may be labeled and the assay may be carried out where the duplex is bound to a surface, so that upon the formation of duplex on the surface, the presence of antibody bound to the duplex can be detected.

[0289] Gene expression, alternatively, may be measured by immunological methods, such as immunohistochemical staining of cells or tissue sections and assay of cell culture or body fluids, to quantitate directly the expression of gene product. Antibodies useful for immunohistochemical staining and/or assay of sample fluids may be either monoclonal or polyclonal, and may be prepared in any mammal. Conveniently, the antibodies may be prepared against a native sequence PRO polypeptide or against a synthetic peptide based on the DNA sequences provided herein or against exogenous sequence fused to PRO DNA and encoding a specific antibody epitope.

#### 5. Purification of Polypeptide

[0290] Forms of PRO may be recovered from culture medium or from host cell lysates. If membrane-bound, it can be released from the membrane using a suitable detergent solution (e.g. Triton-X 100) or by enzymatic cleavage. Cells employed in expression of PRO can be disrupted by various physical or chemical means, such as freeze-thaw cycling, sonication, mechanical disruption, or cell lysing agents.

[0291] It may be desired to purify PRO from recombinant cell proteins or polypeptides. The following procedures are exemplary of suitable purification procedures: by fractionation on an ion-exchange column; ethanol precipitation; reverse phase HPLC; chromatography on silica or on a cation-exchange resin such as DEAE; chromatofocusing; SDS-PAGE; ammonium sulfate precipitation; gel filtration using, for example, Sephadex

G-75; protein A Sepharose columns to remove contaminants such as IgG; and metal chelating columns to bind epitope-tagged forms of the PRO. Various methods of protein purification may be employed and such methods are known in the art and described for example in Deutscher, Methods in Enzymology, 182 (1990); Scopes, Protein Purification: Principles and Practice, Springer-Verlag, New York (1982). The purification step(s) selected will depend, for example, on the nature of the production process used and the particular PRO produced.

#### E. Uses for PRO

**[0292]** Nucleotide sequences (or their complement) encoding PRO have various applications in the art of molecular biology, including uses as hybridization probes, in chromosome and gene mapping and in the generation of anti-sense RNA and DNA. PRO nucleic acid will also be useful for the preparation of PRO polypeptides by the recombinant techniques described herein.

**[0293]** The full-length native sequence PRO gene, or portions thereof, may be used as hybridization probes for a cDNA library to isolate the full-length PRO cDNA or to isolate still other cDNAs (for instance, those encoding naturally-occurring variants of PRO or PRO from other species) which have a desired sequence identity to the native PRO sequence disclosed herein. Optionally, the length of the probes will be about 20 to about 50 bases. The hybridization probes may be derived from at least partially novel regions of the full length native nucleotide sequence wherein those regions may be determined without undue experimentation or from genomic sequences including promoters, enhancer elements and introns of native sequence PRO. By way of example, a screening method will comprise isolating the coding region of the PRO gene using the known DNA sequence to synthesize a selected probe of about 40 bases. Hybridization probes may be labeled by a variety of labels, including radionucleotides such as  $^{32}\text{P}$  or  $^{35}\text{S}$ , or enzymatic labels such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems. Labeled probes having a sequence complementary to that of the PRO gene of the present invention can be used to screen libraries of human cDNA, genomic DNA or mRNA to determine which members of such libraries the probe hybridizes to. Hybridization techniques are described in further detail in the Examples below.

[0294] Any EST sequences disclosed in the present application may similarly be employed as probes, using the methods disclosed herein.

[0295] Other useful fragments of the PRO nucleic acids include antisense or sense oligonucleotides comprising a single-stranded nucleic acid sequence (either RNA or DNA) capable of binding to target PRO mRNA (sense) or PRO DNA (antisense) sequences. Antisense or sense oligonucleotides, according to the present invention, comprise a fragment of the coding region of PRO DNA. Such a fragment generally comprises at least about 14 nucleotides, preferably from about 14 to 30 nucleotides. The ability to derive an antisense or a sense oligonucleotide, based upon a cDNA sequence encoding a given protein is described in, for example, Stein and Cohen (Cancer Res. 48:2659, 1988) and van der Krol et al. (BioTechniques 6:958, 1988).

[0296] Binding of antisense or sense oligonucleotides to target nucleic acid sequences results in the formation of duplexes that block transcription or translation of the target sequence by one of several means, including enhanced degradation of the duplexes, premature termination of transcription or translation, or by other means. The antisense oligonucleotides thus may be used to block expression of PRO proteins. Antisense or sense oligonucleotides further comprise oligonucleotides having modified sugar-phosphodiester backbones (or other sugar linkages, such as those described in WO 91/06629) and wherein such sugar linkages are resistant to endogenous nucleases. Such oligonucleotides with resistant sugar linkages are stable *in vivo* (i.e., capable of resisting enzymatic degradation) but retain sequence specificity to be able to bind to target nucleotide sequences.

[0297] Other examples of sense or antisense oligonucleotides include those oligonucleotides which are covalently linked to organic moieties, such as those described in WO 90/10048, and other moieties that increases affinity of the oligonucleotide for a target nucleic acid sequence, such as poly-(L-lysine). Further still, intercalating agents, such as ellipticine, and alkylating agents or metal complexes may be attached to sense or antisense oligonucleotides to modify binding specificities of the antisense or sense oligonucleotide for the target nucleotide sequence.

[0298] Antisense or sense oligonucleotides may be introduced into a cell containing the target nucleic acid sequence by any gene transfer method, including, for

example, CaPO<sub>4</sub>-mediated DNA transfection, electroporation, or by using gene transfer vectors such as Epstein-Barr virus. In a preferred procedure, an antisense or sense oligonucleotide is inserted into a suitable retroviral vector. A cell containing the target nucleic acid sequence is contacted with the recombinant retroviral vector, either *in vivo* or *ex vivo*. Suitable retroviral vectors include, but are not limited to, those derived from the murine retrovirus M-MuLV, N2 (a retrovirus derived from M-MuLV), or the double copy vectors designated DCT5A, DCT5B and DCT5C (see WO 90/13641).

[0299] Sense or antisense oligonucleotides also may be introduced into a cell containing the target nucleotide sequence by formation of a conjugate with a ligand binding molecule, as described in WO 91/04753. Suitable ligand binding molecules include, but are not limited to, cell surface receptors, growth factors, other cytokines, or other ligands that bind to cell surface receptors. Preferably, conjugation of the ligand binding molecule does not substantially interfere with the ability of the ligand binding molecule to bind to its corresponding molecule or receptor, or block entry of the sense or antisense oligonucleotide or its conjugated version into the cell.

[0300] Alternatively, a sense or an antisense oligonucleotide may be introduced into a cell containing the target nucleic acid sequence by formation of an oligonucleotide-lipid complex, as described in WO 90/10448. The sense or antisense oligonucleotide-lipid complex is preferably dissociated within the cell by an endogenous lipase.

[0301] Antisense or sense RNA or DNA molecules are generally at least about 5 bases in length, about 10 bases in length, about 15 bases in length, about 20 bases in length, about 25 bases in length, about 30 bases in length, about 35 bases in length, about 40 bases in length, about 45 bases in length, about 50 bases in length, about 55 bases in length, about 60 bases in length, about 65 bases in length, about 70 bases in length, about 75 bases in length, about 80 bases in length, about 85 bases in length, about 90 bases in length, about 95 bases in length, about 100 bases in length, or more.

[0302] The probes may also be employed in PCR techniques to generate a pool of sequences for identification of closely related PRO coding sequences.

[0303] Nucleotide sequences encoding a PRO can also be used to construct hybridization probes for mapping the gene which encodes that PRO and for the genetic analysis



of individuals with genetic disorders. The nucleotide sequences provided herein may be mapped to a chromosome and specific regions of a chromosome using known techniques, such as *in situ* hybridization, linkage analysis against known chromosomal markers, and hybridization screening with libraries.

[0304] When the coding sequences for PRO encode a protein which binds to another protein (example, where the PRO is a receptor), the PRO can be used in assays to identify the other proteins or molecules involved in the binding interaction. By such methods, inhibitors of the receptor/ligand binding interaction can be identified. Proteins involved in such binding interactions can also be used to screen for peptide or small molecule inhibitors or agonists of the binding interaction. Also, the receptor PRO can be used to isolate correlative ligand(s). Screening assays can be designed to find lead compounds that mimic the biological activity of a native PRO or a receptor for PRO. Such screening assays will include assays amenable to high-throughput screening of chemical libraries, making them particularly suitable for identifying small molecule drug candidates. Small molecules contemplated include synthetic organic or inorganic compounds. The assays can be performed in a variety of formats, including protein-protein binding assays, biochemical screening assays, immunoassays and cell based assays, which are well characterized in the art.

[0305] Nucleic acids which encode PRO or its modified forms can also be used to generate either transgenic animals or “knock out” animals which, in turn, are useful in the development and screening of therapeutically useful reagents. A transgenic animal (e.g., a mouse or rat) is an animal having cells that contain a transgene, which transgene was introduced into the animal or an ancestor of the animal at a prenatal, e.g., an embryonic stage. A transgene is a DNA which is integrated into the genome of a cell from which a transgenic animal develops. In one embodiment, cDNA encoding PRO can be used to clone genomic DNA encoding PRO in accordance with established techniques and the genomic sequences used to generate transgenic animals that contain cells which express DNA encoding PRO. Methods for generating transgenic animals, particularly animals such as mice or rats, have become conventional in the art and are described, for example, in U.S. Patent Nos. 4,736,866 and 4,870,009. Typically, particular cells would be targeted for PRO transgene incorporation with tissue-specific enhancers. Transgenic animals that include a copy of a transgene encoding

PRO introduced into the germ line of the animal at an embryonic stage can be used to examine the effect of increased expression of DNA encoding PRO. Such animals can be used as tester animals for reagents thought to confer protection from, for example, pathological conditions associated with its overexpression. In accordance with this facet of the invention, an animal is treated with the reagent and a reduced incidence of the pathological condition, compared to untreated animals bearing the transgene, would indicate a potential therapeutic intervention for the pathological condition.

[0306] Alternatively, non-human homologues of PRO can be used to construct a PRO “knock out” animal which has a defective or altered gene encoding PRO as a result of homologous recombination between the endogenous gene encoding PRO and altered genomic DNA encoding PRO introduced into an embryonic stem cell of the animal. For example, cDNA encoding PRO can be used to clone genomic DNA encoding PRO in accordance with established techniques. A portion of the genomic DNA encoding PRO can be deleted or replaced with another gene, such as a gene encoding a selectable marker which can be used to monitor integration. Typically, several kilobases of unaltered flanking DNA (both at the 5’ and 3’ ends) are included in the vector [see e.g., Thomas and Capecchi, Cell, 51:503 (1987) for a description of homologous recombination vectors]. The vector is introduced into an embryonic stem cell line (e.g., by electroporation) and cells in which the introduced DNA has homologously recombined with the endogenous DNA are selected [see e.g., Li et al., Cell, 69:915 (1992)]. The selected cells are then injected into a blastocyst of an animal (e.g., a mouse or rat) to form aggregation chimeras [see e.g., Bradley, in *Teratocarcinomas and Embryonic Stem Cells: A Practical Approach*, E. J. Robertson, ed. (IRL, Oxford, 1987), pp. 113-152]. A chimeric embryo can then be implanted into a suitable pseudopregnant female foster animal and the embryo brought to term to create a “knock out” animal. Progeny harboring the homologously recombined DNA in their germ cells can be identified by standard techniques and used to breed animals in which all cells of the animal contain the homologously recombined DNA. Knockout animals can be characterized for instance, for their ability to defend against certain pathological conditions and for their development of pathological conditions due to absence of the PRO polypeptide.

[0307] Nucleic acid encoding the PRO polypeptides may also be used in gene therapy. In gene therapy applications, genes are introduced into cells in order to achieve *in vivo* synthesis of a therapeutically effective genetic product, for example for replacement of a defective gene. "Gene therapy" includes both conventional gene therapy where a lasting effect is achieved by a single treatment, and the administration of gene therapeutic agents, which involves the one time or repeated administration of a therapeutically effective DNA or mRNA. Antisense RNAs and DNAs can be used as therapeutic agents for blocking the expression of certain genes *in vivo*. It has already been shown that short antisense oligonucleotides can be imported into cells where they act as inhibitors, despite their low intracellular concentrations caused by their restricted uptake by the cell membrane. (Zamecnik *et al.*, Proc. Natl. Acad. Sci. USA 83:4143-4146 [1986]). The oligonucleotides can be modified to enhance their uptake, e.g. by substituting their negatively charged phosphodiester groups by uncharged groups.

[0308] There are a variety of techniques available for introducing nucleic acids into viable cells. The techniques vary depending upon whether the nucleic acid is transferred into cultured cells *in vitro*, or *in vivo* in the cells of the intended host. Techniques suitable for the transfer of nucleic acid into mammalian cells *in vitro* include the use of liposomes, electroporation, microinjection, cell fusion, DEAE-dextran, the calcium phosphate precipitation method, etc. The currently preferred *in vivo* gene transfer techniques include transfection with viral (typically retroviral) vectors and viral coat protein-liposome mediated transfection (Dzau *et al.*, Trends in Biotechnology 11, 205-210 [1993]). In some situations it is desirable to provide the nucleic acid source with an agent that targets the target cells, such as an antibody specific for a cell surface membrane protein or the target cell, a ligand for a receptor on the target cell, etc. Where liposomes are employed, proteins which bind to a cell surface membrane protein associated with endocytosis may be used for targeting and/or to facilitate uptake, e.g. capsid proteins or fragments thereof tropic for a particular cell type, antibodies for proteins which undergo internalization in cycling, proteins that target intracellular localization and enhance intracellular half-life. The technique of receptor-mediated endocytosis is described, for example, by Wu *et al.*, J. Biol. Chem. 262, 4429-4432 (1987); and Wagner *et al.*, Proc. Natl. Acad. Sci. USA 87, 3410-3414 (1990). For

review of gene marking and gene therapy protocols see Anderson et al., Science 256, 808-813 (1992).

[0309] The PRO polypeptides described herein may also be employed as molecular weight markers for protein electrophoresis purposes and the isolated nucleic acid sequences may be used for recombinantly expressing those markers.

[0310] The nucleic acid molecules encoding the PRO polypeptides or fragments thereof described herein are useful for chromosome identification. In this regard, there exists an ongoing need to identify new chromosome markers, since relatively few chromosome marking reagents, based upon actual sequence data are presently available. Each PRO nucleic acid molecule of the present invention can be used as a chromosome marker.

[0311] The PRO polypeptides and nucleic acid molecules of the present invention may also be used diagnostically for tissue typing, wherein the PRO polypeptides of the present invention may be differentially expressed in one tissue as compared to another, preferably in a diseased tissue as compared to a normal tissue of the same tissue type. PRO nucleic acid molecules will find use for generating probes for PCR, Northern analysis, Southern analysis and Western analysis.

[0312] The PRO polypeptides described herein may also be employed as therapeutic agents. The PRO polypeptides of the present invention can be formulated according to known methods to prepare pharmaceutically useful compositions, whereby the PRO product hereof is combined in admixture with a pharmaceutically acceptable carrier vehicle. Therapeutic formulations are prepared for storage by mixing the active ingredient having the desired degree of purity with optional physiologically acceptable carriers, excipients or stabilizers (Remington's Pharmaceutical Sciences 16th edition, Osol, A. Ed. (1980)), in the form of lyophilized formulations or aqueous solutions. Acceptable carriers, excipients or stabilizers are nontoxic to recipients at the dosages and concentrations employed, and include buffers such as phosphate, citrate and other organic acids; antioxidants including ascorbic acid; low molecular weight (less than about 10 residues) polypeptides; proteins, such as serum albumin, gelatin or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone, amino acids such as glycine, glutamine, asparagine, arginine or lysine; monosaccharides, disaccharides and other carbohydrates including glucose, mannose, or dextrans; chelating

agents such as EDTA; sugar alcohols such as mannitol or sorbitol; salt-forming counterions such as sodium; and/or nonionic surfactants such as TWEEN<sup>TM</sup>, PLURONICS<sup>TM</sup> or PEG.

[0313] The formulations to be used for *in vivo* administration must be sterile. This is readily accomplished by filtration through sterile filtration membranes, prior to or following lyophilization and reconstitution.

[0314] Therapeutic compositions herein generally are placed into a container having a sterile access port, for example, an intravenous solution bag or vial having a stopper pierceable by a hypodermic injection needle.

[0315] The route of administration is in accord with known methods, e.g. injection or infusion by intravenous, intraperitoneal, intracerebral, intramuscular, intraocular, intraarterial or intralesional routes, topical administration, or by sustained release systems.

[0316] Dosages and desired drug concentrations of pharmaceutical compositions of the present invention may vary depending on the particular use envisioned. The determination of the appropriate dosage or route of administration is well within the skill of an ordinary physician. Animal experiments provide reliable guidance for the determination of effective doses for human therapy. Interspecies scaling of effective doses can be performed following the principles laid down by Mordenti, J. and Chappell, W. "The use of interspecies scaling in toxicokinetics" In Toxicokinetics and New Drug Development, Yacobi et al., Eds., Pergamon Press, New York 1989, pp. 42-96.

[0317] When *in vivo* administration of a PRO polypeptide or agonist or antagonist thereof is employed, normal dosage amounts may vary from about 10 ng/kg to up to 100 mg/kg of mammal body weight or more per day, preferably about 1 µg/kg/day to 10 mg/kg/day, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature; see, for example, U.S. Pat. Nos. 4,657,760; 5,206,344; or 5,225,212. It is anticipated that different formulations will be effective for different treatment compounds and different disorders, that administration targeting one organ or tissue, for example, may necessitate delivery in a manner different from that to another organ or tissue.

[0318] Where sustained-release administration of a PRO polypeptide is desired in a formulation with release characteristics suitable for the treatment of any disease or disorder requiring administration of the PRO polypeptide, microencapsulation of the PRO polypeptide

is contemplated. Microencapsulation of recombinant proteins for sustained release has been successfully performed with human growth hormone (rhGH), interferon- (rhIFN- ), interleukin-2, and MN rgp120. Johnson et al., Nat. Med., 2:795-799 (1996); Yasuda, Biomed. Ther., 27:1221-1223 (1993); Hora et al., Bio/Technology, 8:755-758 (1990); Cleland, "Design and Production of Single Immunization Vaccines Using Polylactide Polyglycolide Microsphere Systems," in Vaccine Design: The Subunit and Adjuvant Approach, Powell and Newman, eds, (Plenum Press: New York, 1995), pp. 439-462; WO 97/03692, WO 96/40072, WO 96/07399; and U.S. Pat. No. 5,654,010.

[0319] The sustained-release formulations of these proteins were developed using poly-lactic-coglycolic acid (PLGA) polymer due to its biocompatibility and wide range of biodegradable properties. The degradation products of PLGA, lactic and glycolic acids, can be cleared quickly within the human body. Moreover, the degradability of this polymer can be adjusted from months to years depending on its molecular weight and composition. Lewis, "Controlled release of bioactive agents from lactide/glycolide polymer," in: M. Chasin and R. Langer (Eds.), Biodegradable Polymers as Drug Delivery Systems (Marcel Dekker: New York, 1990), pp. 1-41.

[0320] This invention encompasses methods of screening compounds to identify those that mimic the PRO polypeptide (agonists) or prevent the effect of the PRO polypeptide (antagonists). Screening assays for antagonist drug candidates are designed to identify compounds that bind or complex with the PRO polypeptides encoded by the genes identified herein, or otherwise interfere with the interaction of the encoded polypeptides with other cellular proteins. Such screening assays will include assays amenable to high-throughput screening of chemical libraries, making them particularly suitable for identifying small molecule drug candidates.

[0321] The assays can be performed in a variety of formats, including protein-protein binding assays, biochemical screening assays, immunoassays, and cell-based assays, which are well characterized in the art.

[0322] All assays for antagonists are common in that they call for contacting the drug candidate with a PRO polypeptide encoded by a nucleic acid identified herein under conditions and for a time sufficient to allow these two components to interact.

[0323] In binding assays, the interaction is binding and the complex formed can be isolated or detected in the reaction mixture. In a particular embodiment, the PRO polypeptide encoded by the gene identified herein or the drug candidate is immobilized on a solid phase, e.g., on a microtiter plate, by covalent or non-covalent attachments. Non-covalent attachment generally is accomplished by coating the solid surface with a solution of the PRO polypeptide and drying. Alternatively, an immobilized antibody, e.g., a monoclonal antibody, specific for the PRO polypeptide to be immobilized can be used to anchor it to a solid surface. The assay is performed by adding the non-immobilized component, which may be labeled by a detectable label, to the immobilized component, e.g., the coated surface containing the anchored component. When the reaction is complete, the non-reacted components are removed, e.g., by washing, and complexes anchored on the solid surface are detected. When the originally non-immobilized component carries a detectable label, the detection of label immobilized on the surface indicates that complexing occurred. Where the originally non-immobilized component does not carry a label, complexing can be detected, for example, by using a labeled antibody specifically binding the immobilized complex.

[0324] If the candidate compound interacts with but does not bind to a particular PRO polypeptide encoded by a gene identified herein, its interaction with that polypeptide can be assayed by methods well known for detecting protein-protein interactions. Such assays include traditional approaches, such as, e.g., cross-linking, co-immunoprecipitation, and co-purification through gradients or chromatographic columns. In addition, protein-protein interactions can be monitored by using a yeast-based genetic system described by Fields and co-workers (Fields and Song, Nature (London), 340:245-246 (1989); Chien et al., Proc. Natl. Acad. Sci. USA, 88:9578-9582 (1991)) as disclosed by Chevray and Nathans, Proc. Natl. Acad. Sci. USA, 89: 5789-5793 (1991). Many transcriptional activators, such as yeast GAL4, consist of two physically discrete modular domains, one acting as the DNA-binding domain, the other one functioning as the transcription-activation domain. The yeast expression system described in the foregoing publications (generally referred to as the “two-hybrid system”) takes advantage of this property, and employs two hybrid proteins, one in which the target protein is fused to the DNA-binding domain of GAL4, and another, in which candidate activating proteins are fused to the activation domain. The expression of a GAL1-*lacZ* reporter gene

under control of a GAL4-activated promoter depends on reconstitution of GAL4 activity via protein-protein interaction. Colonies containing interacting polypeptides are detected with a chromogenic substrate for  $\beta$ -galactosidase. A complete kit (MATCHMAKER™) for identifying protein-protein interactions between two specific proteins using the two-hybrid technique is commercially available from Clontech. This system can also be extended to map protein domains involved in specific protein interactions as well as to pinpoint amino acid residues that are crucial for these interactions.

[0325] Compounds that interfere with the interaction of a gene encoding a PRO polypeptide identified herein and other intra- or extracellular components can be tested as follows: usually a reaction mixture is prepared containing the product of the gene and the intra- or extracellular component under conditions and for a time allowing for the interaction and binding of the two products. To test the ability of a candidate compound to inhibit binding, the reaction is run in the absence and in the presence of the test compound. In addition, a placebo may be added to a third reaction mixture, to serve as positive control. The binding (complex formation) between the test compound and the intra- or extracellular component present in the mixture is monitored as described hereinabove. The formation of a complex in the control reaction(s) but not in the reaction mixture containing the test compound indicates that the test compound interferes with the interaction of the test compound and its reaction partner.

[0326] To assay for antagonists, the PRO polypeptide may be added to a cell along with the compound to be screened for a particular activity and the ability of the compound to inhibit the activity of interest in the presence of the PRO polypeptide indicates that the compound is an antagonist to the PRO polypeptide. Alternatively, antagonists may be detected by combining the PRO polypeptide and a potential antagonist with membrane-bound PRO polypeptide receptors or recombinant receptors under appropriate conditions for a competitive inhibition assay. The PRO polypeptide can be labeled, such as by radioactivity, such that the number of PRO polypeptide molecules bound to the receptor can be used to determine the effectiveness of the potential antagonist. The gene encoding the receptor can be identified by numerous methods known to those of skill in the art, for example, ligand panning and FACS sorting. Coligan et al., Current Protocols in Immun., 1(2): Chapter 5 (1991). Preferably, expression cloning is employed wherein polyadenylated RNA is prepared from a cell



responsive to the PRO polypeptide and a cDNA library created from this RNA is divided into pools and used to transfect COS cells or other cells that are not responsive to the PRO polypeptide. Transfected cells that are grown on glass slides are exposed to labeled PRO polypeptide. The PRO polypeptide can be labeled by a variety of means including iodination or inclusion of a recognition site for a site-specific protein kinase. Following fixation and incubation, the slides are subjected to autoradiographic analysis. Positive pools are identified and sub-pools are prepared and re-transfected using an interactive sub-pooling and re-screening process, eventually yielding a single clone that encodes the putative receptor.

[0327] As an alternative approach for receptor identification, labeled PRO polypeptide can be photoaffinity-linked with cell membrane or extract preparations that express the receptor molecule. Cross-linked material is resolved by PAGE and exposed to X-ray film. The labeled complex containing the receptor can be excised, resolved into peptide fragments, and subjected to protein micro-sequencing. The amino acid sequence obtained from micro-sequencing would be used to design a set of degenerate oligonucleotide probes to screen a cDNA library to identify the gene encoding the putative receptor.

[0328] In another assay for antagonists, mammalian cells or a membrane preparation expressing the receptor would be incubated with labeled PRO polypeptide in the presence of the candidate compound. The ability of the compound to enhance or block this interaction could then be measured.

[0329] More specific examples of potential antagonists include an oligonucleotide that binds to the fusions of immunoglobulin with PRO polypeptide, and, in particular, antibodies including, without limitation, poly- and monoclonal antibodies and antibody fragments, single-chain antibodies, anti-idiotypic antibodies, and chimeric or humanized versions of such antibodies or fragments, as well as human antibodies and antibody fragments. Alternatively, a potential antagonist may be a closely related protein, for example, a mutated form of the PRO polypeptide that recognizes the receptor but imparts no effect, thereby competitively inhibiting the action of the PRO polypeptide.

[0330] Another potential PRO polypeptide antagonist is an antisense RNA or DNA construct prepared using antisense technology, where, e.g., an antisense RNA or DNA molecule acts to block directly the translation of mRNA by hybridizing to targeted mRNA and

preventing protein translation. Antisense technology can be used to control gene expression through triple-helix formation or antisense DNA or RNA, both of which methods are based on binding of a polynucleotide to DNA or RNA. For example, the 5' coding portion of the polynucleotide sequence, which encodes the mature PRO polypeptides herein, is used to design an antisense RNA oligonucleotide of from about 10 to 40 base pairs in length. A DNA oligonucleotide is designed to be complementary to a region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res., 6:3073 (1979); Cooney et al., Science, 241: 456 (1988); Dervan et al., Science, 251:1360 (1991)), thereby preventing transcription and the production of the PRO polypeptide. The antisense RNA oligonucleotide hybridizes to the mRNA *in vivo* and blocks translation of the mRNA molecule into the PRO polypeptide (antisense - Okano, Neurochem., 56:560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression (CRC Press: Boca Raton, FL, 1988). The oligonucleotides described above can also be delivered to cells such that the antisense RNA or DNA may be expressed *in vivo* to inhibit production of the PRO polypeptide. When antisense DNA is used, oligodeoxyribonucleotides derived from the translation-initiation site, e.g., between about -10 and +10 positions of the target gene nucleotide sequence, are preferred.

[0331] Potential antagonists include small molecules that bind to the active site, the receptor binding site, or growth factor or other relevant binding site of the PRO polypeptide, thereby blocking the normal biological activity of the PRO polypeptide. Examples of small molecules include, but are not limited to, small peptides or peptide-like molecules, preferably soluble peptides, and synthetic non-peptidyl organic or inorganic compounds.

[0332] Ribozymes are enzymatic RNA molecules capable of catalyzing the specific cleavage of RNA. Ribozymes act by sequence-specific hybridization to the complementary target RNA, followed by endonucleolytic cleavage. Specific ribozyme cleavage sites within a potential RNA target can be identified by known techniques. For further details see, e.g., Rossi, Current Biology, 4:469-471 (1994), and PCT publication No. WO 97/33551 (published September 18, 1997).

[0333] Nucleic acid molecules in triple-helix formation used to inhibit transcription should be single-stranded and composed of deoxynucleotides. The base composition of these oligonucleotides is designed such that it promotes triple-helix formation

via Hoogsteen base-pairing rules, which generally require sizeable stretches of purines or pyrimidines on one strand of a duplex. For further details see, e.g., PCT publication No. WO 97/33551, *supra*.

[0334] These small molecules can be identified by any one or more of the screening assays discussed hereinabove and/or by any other screening techniques well known for those skilled in the art.

[0335] Diagnostic and therapeutic uses of the herein disclosed molecules may also be based upon the positive functional assay hits disclosed and described below.

#### F. Anti-PRO Antibodies

[0336] The present invention further provides anti-PRO antibodies. Exemplary antibodies include polyclonal, monoclonal, humanized, bispecific, and heteroconjugate antibodies.

##### 1. Polyclonal Antibodies

[0337] The anti-PRO antibodies may comprise polyclonal antibodies. Methods of preparing polyclonal antibodies are known to the skilled artisan. Polyclonal antibodies can be raised in a mammal, for example, by one or more injections of an immunizing agent and, if desired, an adjuvant. Typically, the immunizing agent and/or adjuvant will be injected in the mammal by multiple subcutaneous or intraperitoneal injections. The immunizing agent may include the PRO polypeptide or a fusion protein thereof. It may be useful to conjugate the immunizing agent to a protein known to be immunogenic in the mammal being immunized. Examples of such immunogenic proteins include but are not limited to keyhole limpet hemocyanin, serum albumin, bovine thyroglobulin, and soybean trypsin inhibitor. Examples of adjuvants which may be employed include Freund's complete adjuvant and MPL-TDM adjuvant (monophosphoryl Lipid A, synthetic trehalose dicorynomycolate). The immunization protocol may be selected by one skilled in the art without undue experimentation.

## 2. Monoclonal Antibodies

[0338] The anti-PRO antibodies may, alternatively, be monoclonal antibodies. Monoclonal antibodies may be prepared using hybridoma methods, such as those described by Kohler and Milstein, Nature, 256:495 (1975). In a hybridoma method, a mouse, hamster, or other appropriate host animal, is typically immunized with an immunizing agent to elicit lymphocytes that produce or are capable of producing antibodies that will specifically bind to the immunizing agent. Alternatively, the lymphocytes may be immunized *in vitro*.

[0339] The immunizing agent will typically include the PRO polypeptide or a fusion protein thereof. Generally, either peripheral blood lymphocytes ("PBLs") are used if cells of human origin are desired, or spleen cells or lymph node cells are used if non-human mammalian sources are desired. The lymphocytes are then fused with an immortalized cell line using a suitable fusing agent, such as polyethylene glycol, to form a hybridoma cell [Goding, Monoclonal Antibodies: Principles and Practice, Academic Press, (1986) pp. 59-103]. Immortalized cell lines are usually transformed mammalian cells, particularly myeloma cells of rodent, bovine and human origin. Usually, rat or mouse myeloma cell lines are employed. The hybridoma cells may be cultured in a suitable culture medium that preferably contains one or more substances that inhibit the growth or survival of the unfused, immortalized cells. For example, if the parental cells lack the enzyme hypoxanthine guanine phosphoribosyl transferase (HGPRT or HPRT), the culture medium for the hybridomas typically will include hypoxanthine, aminopterin, and thymidine ("HAT medium"), which substances prevent the growth of HGPRT-deficient cells.

[0340] Preferred immortalized cell lines are those that fuse efficiently, support stable high level expression of antibody by the selected antibody-producing cells, and are sensitive to a medium such as HAT medium. More preferred immortalized cell lines are murine myeloma lines, which can be obtained, for instance, from the Salk Institute Cell Distribution Center, San Diego, California and the American Type Culture Collection, Manassas, Virginia. Human myeloma and mouse-human heteromyeloma cell lines also have been described for the production of human monoclonal antibodies [Kozbor, J. Immunol., 133:3001 (1984); Brodeur et al., Monoclonal Antibody Production Techniques and Applications, Marcel Dekker, Inc., New York, (1987) pp. 51-63].

[0341] The culture medium in which the hybridoma cells are cultured can then be assayed for the presence of monoclonal antibodies directed against PRO. Preferably, the binding specificity of monoclonal antibodies produced by the hybridoma cells is determined by immunoprecipitation or by an *in vitro* binding assay, such as radioimmunoassay (RIA) or enzyme-linked immunoabsorbent assay (ELISA). Such techniques and assays are known in the art. The binding affinity of the monoclonal antibody can, for example, be determined by the Scatchard analysis of Munson and Pollard, Anal. Biochem., 107:220 (1980).

[0342] After the desired hybridoma cells are identified, the clones may be subcloned by limiting dilution procedures and grown by standard methods [Goding, supra]. Suitable culture media for this purpose include, for example, Dulbecco's Modified Eagle's Medium and RPMI-1640 medium. Alternatively, the hybridoma cells may be grown *in vivo* as ascites in a mammal.

[0343] The monoclonal antibodies secreted by the subclones may be isolated or purified from the culture medium or ascites fluid by conventional immunoglobulin purification procedures such as, for example, protein A-Sepharose, hydroxylapatite chromatography, gel electrophoresis, dialysis, or affinity chromatography.

[0344] The monoclonal antibodies may also be made by recombinant DNA methods, such as those described in U.S. Patent No. 4,816,567. DNA encoding the monoclonal antibodies of the invention can be readily isolated and sequenced using conventional procedures (e.g., by using oligonucleotide probes that are capable of binding specifically to genes encoding the heavy and light chains of murine antibodies). The hybridoma cells of the invention serve as a preferred source of such DNA. Once isolated, the DNA may be placed into expression vectors, which are then transfected into host cells such as simian COS cells, Chinese hamster ovary (CHO) cells, or myeloma cells that do not otherwise produce immunoglobulin protein, to obtain the synthesis of monoclonal antibodies in the recombinant host cells. The DNA also may be modified, for example, by substituting the coding sequence for human heavy and light chain constant domains in place of the homologous murine sequences [U.S. Patent No. 4,816,567; Morrison et al., supra] or by covalently joining to the immunoglobulin coding sequence all or part of the coding sequence for a non-immunoglobulin polypeptide. Such a non-immunoglobulin polypeptide can be

substituted for the constant domains of an antibody of the invention, or can be substituted for the variable domains of one antigen-combining site of an antibody of the invention to create a chimeric bivalent antibody.

[0345] The antibodies may be monovalent antibodies. Methods for preparing monovalent antibodies are well known in the art. For example, one method involves recombinant expression of immunoglobulin light chain and modified heavy chain. The heavy chain is truncated generally at any point in the Fc region so as to prevent heavy chain crosslinking. Alternatively, the relevant cysteine residues are substituted with another amino acid residue or are deleted so as to prevent crosslinking.

[0346] *In vitro* methods are also suitable for preparing monovalent antibodies. Digestion of antibodies to produce fragments thereof, particularly, Fab fragments, can be accomplished using routine techniques known in the art.

### 3. Human and Humanized Antibodies

[0347] The anti-PRO antibodies of the invention may further comprise humanized antibodies or human antibodies. Humanized forms of non-human (e.g., murine) antibodies are chimeric immunoglobulins, immunoglobulin chains or fragments thereof (such as Fv, Fab, Fab', F(ab')<sub>2</sub> or other antigen-binding subsequences of antibodies) which contain minimal sequence derived from non-human immunoglobulin. Humanized antibodies include human immunoglobulins (recipient antibody) in which residues from a complementary determining region (CDR) of the recipient are replaced by residues from a CDR of a non-human species (donor antibody) such as mouse, rat or rabbit having the desired specificity, affinity and capacity. In some instances, Fv framework residues of the human immunoglobulin are replaced by corresponding non-human residues. Humanized antibodies may also comprise residues which are found neither in the recipient antibody nor in the imported CDR or framework sequences. In general, the humanized antibody will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the CDR regions correspond to those of a non-human immunoglobulin and all or substantially all of the FR regions are those of a human immunoglobulin consensus sequence. The humanized antibody optimally also will comprise at least a portion of an immunoglobulin constant region

(Fc), typically that of a human immunoglobulin [Jones et al., Nature, 321:522-525 (1986); Riechmann et al., Nature, 332:323-329 (1988); and Presta, Curr. Op. Struct. Biol., 2:593-596 (1992)].

[0348] Methods for humanizing non-human antibodies are well known in the art. Generally, a humanized antibody has one or more amino acid residues introduced into it from a source which is non-human. These non-human amino acid residues are often referred to as “import” residues, which are typically taken from an “import” variable domain. Humanization can be essentially performed following the method of Winter and co-workers [Jones et al., Nature, 321:522-525 (1986); Riechmann et al., Nature, 332:323-327 (1988); Verhoeven et al., Science, 239:1534-1536 (1988)], by substituting rodent CDRs or CDR sequences for the corresponding sequences of a human antibody. Accordingly, such “humanized” antibodies are chimeric antibodies (U.S. Patent No. 4,816,567), wherein substantially less than an intact human variable domain has been substituted by the corresponding sequence from a non-human species. In practice, humanized antibodies are typically human antibodies in which some CDR residues and possibly some FR residues are substituted by residues from analogous sites in rodent antibodies.

[0349] Human antibodies can also be produced using various techniques known in the art, including phage display libraries [Hoogenboom and Winter, J. Mol. Biol., 227:381 (1991); Marks et al., J. Mol. Biol., 222:581 (1991)]. The techniques of Cole et al. and Boerner et al. are also available for the preparation of human monoclonal antibodies (Cole et al., Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, p. 77 (1985) and Boerner et al., J. Immunol., 147(1):86-95 (1991)]. Similarly, human antibodies can be made by introducing of human immunoglobulin loci into transgenic animals, e.g., mice in which the endogenous immunoglobulin genes have been partially or completely inactivated. Upon challenge, human antibody production is observed, which closely resembles that seen in humans in all respects, including gene rearrangement, assembly, and antibody repertoire. This approach is described, for example, in U.S. Patent Nos. 5,545,807; 5,545,806; 5,569,825; 5,625,126; 5,633,425; 5,661,016, and in the following scientific publications: Marks *et al.*, Bio/Technology 10, 779-783 (1992); Lonberg *et al.*, Nature 368 856-859 (1994); Morrison, Nature 368, 812-13

(1994); Fishwild *et al.*, Nature Biotechnology **14**, 845-51 (1996); Neuberger, Nature Biotechnology **14**, 826 (1996); Lonberg and Huszar, Intern. Rev. Immunol. **13** 65-93 (1995).

[0350] The antibodies may also be affinity matured using known selection and/or mutagenesis methods as described above. Preferred affinity matured antibodies have an affinity which is five times, more preferably 10 times, even more preferably 20 or 30 times greater than the starting antibody (generally murine, humanized or human) from which the matured antibody is prepared.

#### 4. Bispecific Antibodies

[0351] Bispecific antibodies are monoclonal, preferably human or humanized, antibodies that have binding specificities for at least two different antigens. In the present case, one of the binding specificities is for the PRO, the other one is for any other antigen, and preferably for a cell-surface protein or receptor or receptor subunit.

[0352] Methods for making bispecific antibodies are known in the art. Traditionally, the recombinant production of bispecific antibodies is based on the co-expression of two immunoglobulin heavy-chain/light-chain pairs, where the two heavy chains have different specificities [Milstein and Cuello, Nature, **305**:537-539 (1983)]. Because of the random assortment of immunoglobulin heavy and light chains, these hybridomas (quadromas) produce a potential mixture of ten different antibody molecules, of which only one has the correct bispecific structure. The purification of the correct molecule is usually accomplished by affinity chromatography steps. Similar procedures are disclosed in WO 93/08829, published 13 May 1993, and in Traunecker *et al.*, EMBO J., **10**:3655-3659 (1991).

[0353] Antibody variable domains with the desired binding specificities (antibody-antigen combining sites) can be fused to immunoglobulin constant domain sequences. The fusion preferably is with an immunoglobulin heavy-chain constant domain, comprising at least part of the hinge, CH2, and CH3 regions. It is preferred to have the first heavy-chain constant region (CH1) containing the site necessary for light-chain binding present in at least one of the fusions. DNAs encoding the immunoglobulin heavy-chain fusions and, if desired, the immunoglobulin light chain, are inserted into separate expression vectors,



and are co-transfected into a suitable host organism. For further details of generating bispecific antibodies see, for example, Suresh et al., Methods in Enzymology, 121:210 (1986).

[0354] According to another approach described in WO 96/27011, the interface between a pair of antibody molecules can be engineered to maximize the percentage of heterodimers which are recovered from recombinant cell culture. The preferred interface comprises at least a part of the CH3 region of an antibody constant domain. In this method, one or more small amino acid side chains from the interface of the first antibody molecule are replaced with larger side chains (e.g. tyrosine or tryptophan). Compensatory “cavities” of identical or similar size to the large side chain(s) are created on the interface of the second antibody molecule by replacing large amino acid side chains with smaller ones (e.g. alanine or threonine). This provides a mechanism for increasing the yield of the heterodimer over other unwanted end-products such as homodimers.

[0355] Bispecific antibodies can be prepared as full length antibodies or antibody fragments (e.g. F(ab')<sub>2</sub> bispecific antibodies). Techniques for generating bispecific antibodies from antibody fragments have been described in the literature. For example, bispecific antibodies can be prepared using chemical linkage. Brennan *et al.*, Science 229:81 (1985) describe a procedure wherein intact antibodies are proteolytically cleaved to generate F(ab')<sub>2</sub> fragments. These fragments are reduced in the presence of the dithiol complexing agent sodium arsenite to stabilize vicinal dithiols and prevent intermolecular disulfide formation. The Fab' fragments generated are then converted to thionitrobenzoate (TNB) derivatives. One of the Fab'-TNB derivatives is then reconverted to the Fab'-thiol by reduction with mercaptoethylamine and is mixed with an equimolar amount of the other Fab'-TNB derivative to form the bispecific antibody. The bispecific antibodies produced can be used as agents for the selective immobilization of enzymes.

[0356] Fab' fragments may be directly recovered from *E. coli* and chemically coupled to form bispecific antibodies. Shalaby *et al.*, J. Exp. Med. 175:217-225 (1992) describe the production of a fully humanized bispecific antibody F(ab')<sub>2</sub> molecule. Each Fab' fragment was separately secreted from *E. coli* and subjected to directed chemical coupling *in vitro* to form the bispecific antibody. The bispecific antibody thus formed was able to bind to

cells overexpressing the ErbB2 receptor and normal human T cells, as well as trigger the lytic activity of human cytotoxic lymphocytes against human breast tumor targets.

[0357] Various technique for making and isolating bispecific antibody fragments directly from recombinant cell culture have also been described. For example, bispecific antibodies have been produced using leucine zippers. Kostelny *et al.*, J. Immunol. 148(5):1547-1553 (1992). The leucine zipper peptides from the Fos and Jun proteins were linked to the Fab' portions of two different antibodies by gene fusion. The antibody homodimers were reduced at the hinge region to form monomers and then re-oxidized to form the antibody heterodimers. This method can also be utilized for the production of antibody homodimers. The "diabody" technology described by Hollinger *et al.*, Proc. Natl. Acad. Sci. USA 90:6444-6448 (1993) has provided an alternative mechanism for making bispecific antibody fragments. The fragments comprise a heavy-chain variable domain (V<sub>H</sub>) connected to a light-chain variable domain (V<sub>L</sub>) by a linker which is too short to allow pairing between the two domains on the same chain. Accordingly, the V<sub>H</sub> and V<sub>L</sub> domains of one fragment are forced to pair with the complementary V<sub>L</sub> and V<sub>H</sub> domains of another fragment, thereby forming two antigen-binding sites. Another strategy for making bispecific antibody fragments by the use of single-chain Fv (sFv) dimers has also been reported. See, Gruber *et al.*, J. Immunol. 152:5368 (1994).

[0358] Antibodies with more than two valencies are contemplated. For example, trispecific antibodies can be prepared. Tutt *et al.*, J. Immunol. 147:60 (1991).

[0359] Exemplary bispecific antibodies may bind to two different epitopes on a given PRO polypeptide herein. Alternatively, an anti-PRO polypeptide arm may be combined with an arm which binds to a triggering molecule on a leukocyte such as a T-cell receptor molecule (e.g. CD2, CD3, CD28, or B7), or Fc receptors for IgG (Fc $\gamma$ R), such as Fc $\gamma$ RI (CD64), Fc $\gamma$ RII (CD32) and Fc $\gamma$ RIII (CD16) so as to focus cellular defense mechanisms to the cell expressing the particular PRO polypeptide. Bispecific antibodies may also be used to localize cytotoxic agents to cells which express a particular PRO polypeptide. These antibodies possess a PRO-binding arm and an arm which binds a cytotoxic agent or a radionuclide chelator, such as EOTUBE, DPTA, DOTA, or TETA. Another bispecific antibody of interest binds the PRO polypeptide and further binds tissue factor (TF).

## 5. Heteroconjugate Antibodies

[0360] Heteroconjugate antibodies are also within the scope of the present invention. Heteroconjugate antibodies are composed of two covalently joined antibodies. Such antibodies have, for example, been proposed to target immune system cells to unwanted cells [U.S. Patent No. 4,676,980], and for treatment of HIV infection [WO 91/00360; WO 92/200373; EP 03089]. It is contemplated that the antibodies may be prepared *in vitro* using known methods in synthetic protein chemistry, including those involving crosslinking agents. For example, immunotoxins may be constructed using a disulfide exchange reaction or by forming a thioether bond. Examples of suitable reagents for this purpose include iminothiolate and methyl-4-mercaptobutyrimidate and those disclosed, for example, in U.S. Patent No. 4,676,980.

## 6. Effector Function Engineering

[0361] It may be desirable to modify the antibody of the invention with respect to effector function, so as to enhance, *e.g.*, the effectiveness of the antibody in treating cancer. For example, cysteine residue(s) may be introduced into the Fc region, thereby allowing interchain disulfide bond formation in this region. The homodimeric antibody thus generated may have improved internalization capability and/or increased complement-mediated cell killing and antibody-dependent cellular cytotoxicity (ADCC). See Caron *et al.*, J. Exp Med., **176**: 1191-1195 (1992) and Shopes, J. Immunol., **148**: 2918-2922 (1992). Homodimeric antibodies with enhanced anti-tumor activity may also be prepared using heterobifunctional cross-linkers as described in Wolff *et al.* Cancer Research, **53**: 2560-2565 (1993). Alternatively, an antibody can be engineered that has dual Fc regions and may thereby have enhanced complement lysis and ADCC capabilities. See Stevenson *et al.*, Anti-Cancer Drug Design, **3**: 219-230 (1989).

## 7. Immunoconjugates

[0362] The invention also pertains to immunoconjugates comprising an antibody conjugated to a cytotoxic agent such as a chemotherapeutic agent, toxin (*e.g.*, an enzymatically

active toxin of bacterial, fungal, plant, or animal origin, or fragments thereof), or a radioactive isotope (*i.e.*, a radioconjugate).

[0363] Chemotherapeutic agents useful in the generation of such immunoconjugates have been described above. Enzymatically active toxins and fragments thereof that can be used include diphtheria A chain, nonbinding active fragments of diphtheria toxin, exotoxin A chain (from *Pseudomonas aeruginosa*), ricin A chain, abrin A chain, modeccin A chain, alpha-sarcin, *Aleurites fordii* proteins, dianthin proteins, *Phytolaca americana* proteins (PAPI, PAPII, and PAP-S), momordica charantia inhibitor, curcin, crotin, sapaonaria officinalis inhibitor, gelonin, mitogellin, restrictocin, phenomycin, enomycin, and the tricothecenes. A variety of radionuclides are available for the production of radioconjugated antibodies. Examples include  $^{212}\text{Bi}$ ,  $^{131}\text{I}$ ,  $^{131}\text{In}$ ,  $^{90}\text{Y}$ , and  $^{186}\text{Re}$ .

[0364] Conjugates of the antibody and cytotoxic agent are made using a variety of bifunctional protein-coupling agents such as N-succinimidyl-3-(2-pyridyldithiol) propionate (SPDP), iminothiolane (IT), bifunctional derivatives of imidoesters (such as dimethyl adipimidate HCL), active esters (such as disuccinimidyl suberate), aldehydes (such as glutaraldehyde), bis-azido compounds (such as bis (p-azidobenzoyl) hexanediamine), bis-diazonium derivatives (such as bis-(p-diazoniumbenzoyl)-ethylenediamine), diisocyanates (such as tolyene 2,6-diisocyanate), and bis-active fluorine compounds (such as 1,5-difluoro-2,4-dinitrobenzene). For example, a ricin immunotoxin can be prepared as described in Vitetta *et al.*, Science, 238: 1098 (1987). Carbon-14-labeled 1-isothiocyanatobenzyl-3-methyldiethylene triaminepentaacetic acid (MX-DTPA) is an exemplary chelating agent for conjugation of radionucleotide to the antibody. See WO94/11026.

[0365] In another embodiment, the antibody may be conjugated to a “receptor” (such streptavidin) for utilization in tumor pretargeting wherein the antibody-receptor conjugate is administered to the patient, followed by removal of unbound conjugate from the circulation using a clearing agent and then administration of a “ligand” (*e.g.*, avidin) that is conjugated to a cytotoxic agent (*e.g.*, a radionucleotide).

## 8. Immunoliposomes

[0366] The antibodies disclosed herein may also be formulated as immunoliposomes. Liposomes containing the antibody are prepared by methods known in the art, such as described in Epstein *et al.*, Proc. Natl. Acad. Sci. USA, 82: 3688 (1985); Hwang *et al.*, Proc. Natl. Acad. Sci. USA, 77: 4030 (1980); and U.S. Pat. Nos. 4,485,045 and 4,544,545. Liposomes with enhanced circulation time are disclosed in U.S. Patent No. 5,013,556.

[0367] Particularly useful liposomes can be generated by the reverse-phase evaporation method with a lipid composition comprising phosphatidylcholine, cholesterol, and PEG-derivatized phosphatidylethanolamine (PEG-PE). Liposomes are extruded through filters of defined pore size to yield liposomes with the desired diameter. Fab' fragments of the antibody of the present invention can be conjugated to the liposomes as described in Martin *et al.*, J. Biol. Chem., 257: 286-288 (1982) via a disulfide-interchange reaction. A chemotherapeutic agent (such as Doxorubicin) is optionally contained within the liposome. See Gabizon *et al.*, J. National Cancer Inst., 81(19): 1484 (1989).

## 9. Pharmaceutical Compositions of Antibodies

[0368] Antibodies specifically binding a PRO polypeptide identified herein, as well as other molecules identified by the screening assays disclosed hereinbefore, can be administered for the treatment of various disorders in the form of pharmaceutical compositions.

[0369] If the PRO polypeptide is intracellular and whole antibodies are used as inhibitors, internalizing antibodies are preferred. However, lipofections or liposomes can also be used to deliver the antibody, or an antibody fragment, into cells. Where antibody fragments are used, the smallest inhibitory fragment that specifically binds to the binding domain of the target protein is preferred. For example, based upon the variable-region sequences of an antibody, peptide molecules can be designed that retain the ability to bind the target protein sequence. Such peptides can be synthesized chemically and/or produced by recombinant DNA technology. See, *e.g.*, Marasco *et al.*, Proc. Natl. Acad. Sci. USA, 90: 7889-7893 (1993).

[0370] The formulation herein may also contain more than one active compound as necessary for the particular indication being treated, preferably those with complementary

activities that do not adversely affect each other. Alternatively, or in addition, the composition may comprise an agent that enhances its function, such as, for example, a cytotoxic agent, cytokine, chemotherapeutic agent, or growth-inhibitory agent. Such molecules are suitably present in combination in amounts that are effective for the purpose intended.

[0371] The active ingredients may also be entrapped in microcapsules prepared, for example, by coacervation techniques or by interfacial polymerization, for example, hydroxymethylcellulose or gelatin-microcapsules and poly-(methylmethacrylate) microcapsules, respectively, in colloidal drug delivery systems (for example, liposomes, albumin microspheres, microemulsions, nano-particles, and nanocapsules) or in macroemulsions. Such techniques are disclosed in Remington's Pharmaceutical Sciences, *supra*.

[0372] The formulations to be used for *in vivo* administration must be sterile. This is readily accomplished by filtration through sterile filtration membranes.

[0373] Sustained-release preparations may be prepared. Suitable examples of sustained-release preparations include semipermeable matrices of solid hydrophobic polymers containing the antibody, which matrices are in the form of shaped articles, *e.g.*, films, or microcapsules. Examples of sustained-release matrices include polyesters, hydrogels (for example, poly(2-hydroxyethyl-methacrylate), or poly(vinylalcohol)), polylactides (U.S. Pat. No. 3,773,919), copolymers of L-glutamic acid and ? ethyl-L-glutamate, non-degradable ethylene-vinyl acetate, degradable lactic acid-glycolic acid copolymers such as the LUPRON DEPOT<sup>TM</sup> (injectable microspheres composed of lactic acid-glycolic acid copolymer and leuprolide acetate), and poly-D-(-)-3-hydroxybutyric acid. While polymers such as ethylene-vinyl acetate and lactic acid-glycolic acid enable release of molecules for over 100 days, certain hydrogels release proteins for shorter time periods. When encapsulated antibodies remain in the body for a long time, they may denature or aggregate as a result of exposure to moisture at 37°C, resulting in a loss of biological activity and possible changes in immunogenicity. Rational strategies can be devised for stabilization depending on the mechanism involved. For example, if the aggregation mechanism is discovered to be intermolecular S-S bond formation through thio-disulfide interchange, stabilization may be achieved by modifying sulfhydryl residues, lyophilizing from acidic solutions, controlling

moisture content, using appropriate additives, and developing specific polymer matrix compositions.

G. Uses for anti-PRO Antibodies

[0374] The anti-PRO antibodies of the invention have various utilities. For example, anti-PRO antibodies may be used in diagnostic assays for PRO, *e.g.*, detecting its expression (and in some cases, differential expression) in specific cells, tissues, or serum. Various diagnostic assay techniques known in the art may be used, such as competitive binding assays, direct or indirect sandwich assays and immunoprecipitation assays conducted in either heterogeneous or homogeneous phases [Zola, Monoclonal Antibodies: A Manual of Techniques, CRC Press, Inc. (1987) pp. 147-158]. The antibodies used in the diagnostic assays can be labeled with a detectable moiety. The detectable moiety should be capable of producing, either directly or indirectly, a detectable signal. For example, the detectable moiety may be a radioisotope, such as  $^3\text{H}$ ,  $^{14}\text{C}$ ,  $^{32}\text{P}$ ,  $^{35}\text{S}$ , or  $^{125}\text{I}$ , a fluorescent or chemiluminescent compound, such as fluorescein isothiocyanate, rhodamine, or luciferin, or an enzyme, such as alkaline phosphatase, beta-galactosidase or horseradish peroxidase. Any method known in the art for conjugating the antibody to the detectable moiety may be employed, including those methods described by Hunter et al., Nature, 144:945 (1962); David et al., Biochemistry, 13:1014 (1974); Pain et al., J. Immunol. Meth., 40:219 (1981); and Nygren, J. Histochem. and Cytochem., 30:407 (1982).

[0375] Anti-PRO antibodies also are useful for the affinity purification of PRO from recombinant cell culture or natural sources. In this process, the antibodies against PRO are immobilized on a suitable support, such a Sephadex resin or filter paper, using methods well known in the art. The immobilized antibody then is contacted with a sample containing the PRO to be purified, and thereafter the support is washed with a suitable solvent that will remove substantially all the material in the sample except the PRO, which is bound to the immobilized antibody. Finally, the support is washed with another suitable solvent that will release the PRO from the antibody.

[0376] The following examples are offered for illustrative purposes only, and are not intended to limit the scope of the present invention in any way.

[0377] All patent and literature references cited in the present specification are hereby incorporated by reference in their entirety.

### EXAMPLES

[0378] Commercially available reagents referred to in the examples were used according to manufacturer's instructions unless otherwise indicated. The source of those cells identified in the following examples, and throughout the specification, by ATCC accession numbers is the American Type Culture Collection, Manassas, VA.

#### EXAMPLE 1

##### Extracellular Domain Homology Screening to Identify Novel Polypeptides and cDNA

##### Encoding Therefor

[0379] The extracellular domain (ECD) sequences (including the secretion signal sequence, if any) from about 950 known secreted proteins from the Swiss-Prot public database were used to search EST databases. The EST databases included public databases (e.g., Dayhoff, GenBank), and proprietary databases (e.g. LIFESEQ<sup>TM</sup>, Incyte Pharmaceuticals, Palo Alto, CA). The search was performed using the computer program BLAST or BLAST-2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)) as a comparison of the ECD protein sequences to a 6 frame translation of the EST sequences. Those comparisons with a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into consensus DNA sequences with the program "phrap" (Phil Green, University of Washington, Seattle, WA).

[0380] Using this extracellular domain homology screen, consensus DNA sequences were assembled relative to the other identified EST sequences using phrap. In addition, the consensus DNA sequences obtained were often (but not always) extended using repeated cycles of BLAST or BLAST-2 and phrap to extend the consensus sequence as far as possible using the sources of EST sequences discussed above.

[0381] Based upon the consensus sequences obtained as described above, oligonucleotides were then synthesized and used to identify by PCR a cDNA library that contained the sequence of interest and for use as probes to isolate a clone of the full-length



coding sequence for a PRO polypeptide. Forward and reverse PCR primers generally range from 20 to 30 nucleotides and are often designed to give a PCR product of about 100-1000 bp in length. The probe sequences are typically 40-55 bp in length. In some cases, additional oligonucleotides are synthesized when the consensus sequence is greater than about 1-1.5kbp.

In order to screen several libraries for a full-length clone, DNA from the libraries was screened by PCR amplification, as per Ausubel et al., Current Protocols in Molecular Biology, with the PCR primer pair. A positive library was then used to isolate clones encoding the gene of interest using the probe oligonucleotide and one of the primer pairs.

[0382] The cDNA libraries used to isolate the cDNA clones were constructed by standard methods using commercially available reagents such as those from Invitrogen, San Diego, CA. The cDNA was primed with oligo dT containing a NotI site, linked with blunt to SalI hemikinased adaptors, cleaved with NotI, sized appropriately by gel electrophoresis, and cloned in a defined orientation into a suitable cloning vector (such as pRKB or pRKD; pRK5B is a precursor of pRK5D that does not contain the SfiI site; see, Holmes et al., Science, 253:1278-1280 (1991)) in the unique XhoI and NotI sites.

## EXAMPLE 2

### Isolation of cDNA clones by Amylase Screening

#### 1. Preparation of oligo dT primed cDNA library

[0383] mRNA was isolated from a human tissue of interest using reagents and protocols from Invitrogen, San Diego, CA (Fast Track 2). This RNA was used to generate an oligo dT primed cDNA library in the vector pRK5D using reagents and protocols from Life Technologies, Gaithersburg, MD (Super Script Plasmid System). In this procedure, the double stranded cDNA was sized to greater than 1000 bp and the SalI/NotI linkered cDNA was cloned into XhoI/NotI cleaved vector. pRK5D is a cloning vector that has an sp6 transcription initiation site followed by an SfiI restriction enzyme site preceding the XhoI/NotI cDNA cloning sites.

## 2. Preparation of random primed cDNA library

[0384] A secondary cDNA library was generated in order to preferentially represent the 5' ends of the primary cDNA clones. Sp6 RNA was generated from the primary library (described above), and this RNA was used to generate a random primed cDNA library in the vector pSST-AMY.0 using reagents and protocols from Life Technologies (Super Script Plasmid System, referenced above). In this procedure the double stranded cDNA was sized to 500-1000 bp, linked with blunt to NotI adaptors, cleaved with SfiI, and cloned into SfiI/NotI cleaved vector. pSST-AMY.0 is a cloning vector that has a yeast alcohol dehydrogenase promoter preceding the cDNA cloning sites and the mouse amylase sequence (the mature sequence without the secretion signal) followed by the yeast alcohol dehydrogenase terminator, after the cloning sites. Thus, cDNAs cloned into this vector that are fused in frame with amylase sequence will lead to the secretion of amylase from appropriately transfected yeast colonies.

## 3. Transformation and Detection

[0385] DNA from the library described in paragraph 2 above was chilled on ice to which was added electrocompetent DH10B bacteria (Life Technologies, 20 ml). The bacteria and vector mixture was then electroporated as recommended by the manufacturer. Subsequently, SOC media (Life Technologies, 1 ml) was added and the mixture was incubated at 37°C for 30 minutes. The transformants were then plated onto 20 standard 150 mm LB plates containing ampicillin and incubated for 16 hours (37°C). Positive colonies were scraped off the plates and the DNA was isolated from the bacterial pellet using standard protocols, e.g. CsCl-gradient. The purified DNA was then carried on to the yeast protocols below.

[0386] The yeast methods were divided into three categories: (1) Transformation of yeast with the plasmid/cDNA combined vector; (2) Detection and isolation of yeast clones secreting amylase; and (3) PCR amplification of the insert directly from the yeast colony and purification of the DNA for sequencing and further analysis.

[0387] The yeast strain used was HD56-5A (ATCC-90785). This strain has the following genotype: MAT alpha, ura3-52, leu2-3, leu2-112, his3-11, his3-15, MAL<sup>+</sup>, SUC<sup>+</sup>, GAL<sup>+</sup>. Preferably, yeast mutants can be employed that have deficient post-translational

pathways. Such mutants may have translocation deficient alleles in *sec71*, *sec72*, *sec62*, with truncated *sec71* being most preferred. Alternatively, antagonists (including antisense nucleotides and/or ligands) which interfere with the normal operation of these genes, other proteins implicated in this post translation pathway (e.g., SEC61p, SEC72p, SEC62p, SEC63p, TDJ1p or SSA1p-4p) or the complex formation of these proteins may also be preferably employed in combination with the amylase-expressing yeast.

**[0388]** Transformation was performed based on the protocol outlined by Gietz et al., Nucl. Acid. Res., 20:1425 (1992). Transformed cells were then inoculated from agar into YEPD complex media broth (100 ml) and grown overnight at 30°C. The YEPD broth was prepared as described in Kaiser et al., Methods in Yeast Genetics, Cold Spring Harbor Press, Cold Spring Harbor, NY, p. 207 (1994). The overnight culture was then diluted to about  $2 \times 10^6$  cells/ml (approx. OD<sub>600</sub>=0.1) into fresh YEPD broth (500 ml) and regrown to  $1 \times 10^7$  cells/ml (approx. OD<sub>600</sub>=0.4-0.5).

**[0389]** The cells were then harvested and prepared for transformation by transfer into GS3 rotor bottles in a Sorval GS3 rotor at 5,000 rpm for 5 minutes, the supernatant discarded, and then resuspended into sterile water, and centrifuged again in 50 ml falcon tubes at 3,500 rpm in a Beckman GS-6KR centrifuge. The supernatant was discarded and the cells were subsequently washed with LiAc/TE (10 ml, 10 mM Tris-HCl, 1 mM EDTA pH 7.5, 100 mM Li<sub>2</sub>OOCCH<sub>3</sub>), and resuspended into LiAc/TE (2.5 ml).

**[0390]** Transformation took place by mixing the prepared cells (100 µl) with freshly denatured single stranded salmon testes DNA (Lofstrand Labs, Gaithersburg, MD) and transforming DNA (1 µg, vol. < 10 µl) in microfuge tubes. The mixture was mixed briefly by vortexing, then 40% PEG/TE (600 µl, 40% polyethylene glycol-4000, 10 mM Tris-HCl, 1 mM EDTA, 100 mM Li<sub>2</sub>OOCCH<sub>3</sub>, pH 7.5) was added. This mixture was gently mixed and incubated at 30°C while agitating for 30 minutes. The cells were then heat shocked at 42°C for 15 minutes, and the reaction vessel centrifuged in a microfuge at 12,000 rpm for 5-10 seconds, decanted and resuspended into TE (500 µl, 10 mM Tris-HCl, 1 mM EDTA pH 7.5) followed by recentrifugation. The cells were then diluted into TE (1 ml) and aliquots (200 µl) were spread onto the selective media previously prepared in 150 mm growth plates (VWR).

[0391] Alternatively, instead of multiple small reactions, the transformation was performed using a single, large scale reaction, wherein reagent amounts were scaled up accordingly.

[0392] The selective media used was a synthetic complete dextrose agar lacking uracil (SCD-Ura) prepared as described in Kaiser et al., Methods in Yeast Genetics, Cold Spring Harbor Press, Cold Spring Harbor, NY, p. 208-210 (1994). Transformants were grown at 30°C for 2-3 days.

[0393] The detection of colonies secreting amylase was performed by including red starch in the selective growth media. Starch was coupled to the red dye (Reactive Red-120, Sigma) as per the procedure described by Biely et al., Anal. Biochem., 172:176-179 (1988). The coupled starch was incorporated into the SCD-Ura agar plates at a final concentration of 0.15% (w/v), and was buffered with potassium phosphate to a pH of 7.0 (50-100 mM final concentration).

[0394] The positive colonies were picked and streaked across fresh selective media (onto 150 mm plates) in order to obtain well isolated and identifiable single colonies. Well isolated single colonies positive for amylase secretion were detected by direct incorporation of red starch into buffered SCD-Ura agar. Positive colonies were determined by their ability to break down starch resulting in a clear halo around the positive colony visualized directly.

#### 4. Isolation of DNA by PCR Amplification

[0395] When a positive colony was isolated, a portion of it was picked by a toothpick and diluted into sterile water (30 µl) in a 96 well plate. At this time, the positive colonies were either frozen and stored for subsequent analysis or immediately amplified. An aliquot of cells (5 µl) was used as a template for the PCR reaction in a 25 µl volume containing: 0.5 µl K<sub>l</sub>entaq KLENTAQ (a 5'-exo minus N-terminal deletion of Taq DNA polymerase available from Clontech, Palo Alto, CA); 4.0 µl 10 mM dNTP's (Perkin Elmer-Cetus); 2.5 µl Kentaq buffer (Clontech); 0.25 µl forward oligo 1; 0.25 µl reverse oligo 2; 12.5 µl distilled water. The sequence of the forward oligonucleotide 1 was:

5'-TGTAACGACGGCCAGTTAAATAGACCTGCAATTATTAATCT-3' (SEQ ID NO:169)

The sequence of reverse oligonucleotide 2 was:

5'-CAGGAAACAGCTATGACCACCTGCACACCTGCAAATCCATT-3' (SEQ ID NO:170)

[0396] PCR was then performed as follows:

- |    |               |          |                  |
|----|---------------|----------|------------------|
| a. | Denature      | 92°C,    | 5 minutes        |
| b. | 3 cycles of:  | Denature | 92°C, 30 seconds |
|    |               | Anneal   | 59°C, 30 seconds |
|    |               | Extend   | 72°C, 60 seconds |
| c. | 3 cycles of:  | Denature | 92°C, 30 seconds |
|    |               | Anneal   | 57°C, 30 seconds |
|    |               | Extend   | 72°C, 60 seconds |
| d. | 25 cycles of: | Denature | 92°C, 30 seconds |
|    |               | Anneal   | 55°C, 30 seconds |
|    |               | Extend   | 72°C, 60 seconds |
| e. | Hold          | 4°C      |                  |

[0397] The underlined regions of the oligonucleotides annealed to the ADH promoter region and the amylase region, respectively, and amplified a 307 bp region from vector pSST-AMY.0 when no insert was present. Typically, the first 18 nucleotides of the 5' end of these oligonucleotides contained annealing sites for the sequencing primers. Thus, the total product of the PCR reaction from an empty vector was 343 bp. However, signal sequence-fused cDNA resulted in considerably longer nucleotide sequences.

[0398] Following the PCR, an aliquot of the reaction (5 µl) was examined by agarose gel electrophoresis in a 1% agarose gel using a Tris-Borate-EDTA (TBE) buffering system as described by Sambrook et al., supra. Clones resulting in a single strong PCR product larger than 400 bp were further analyzed by DNA sequencing after purification with a 96 ~~Qiaquick~~ QIAQUICK PCR clean-up column (Qiagen Inc., Chatsworth, CA).

### EXAMPLE 3

#### Isolation of cDNA Clones Using Signal Algorithm Analysis

[0399] Various polypeptide-encoding nucleic acid sequences were identified by applying a proprietary signal sequence finding algorithm developed by Genentech, Inc. (South San Francisco, CA) upon ESTs as well as clustered and assembled EST fragments from public (e.g., GenBank) and/or private (LIFESEQ®, Incyte Pharmaceuticals, Inc., Palo Alto, CA) databases. The signal sequence algorithm computes a secretion signal score based on the character of the DNA nucleotides surrounding the first and optionally the second methionine codon(s) (ATG) at the 5'-end of the sequence or sequence fragment under consideration. The nucleotides following the first ATG must code for at least 35 unambiguous amino acids without any stop codons. If the first ATG has the required amino acids, the second is not examined. If neither meets the requirement, the candidate sequence is not scored. In order to determine whether the EST sequence contains an authentic signal sequence, the DNA and corresponding amino acid sequences surrounding the ATG codon are scored using a set of seven sensors (evaluation parameters) known to be associated with secretion signals. Use of this algorithm resulted in the identification of numerous polypeptide-encoding nucleic acid sequences.

### EXAMPLE 4

#### Isolation of cDNA Clones Encoding Human PRO Polypeptides

[0400] Using the techniques described in Examples 1 to 3 above, numerous full-length cDNA clones were identified as encoding PRO polypeptides as disclosed herein. These cDNAs were then deposited under the terms of the Budapest Treaty with the American Type Culture Collection, 10801 University Blvd., Manassas, VA 20110-2209, USA (ATCC) as shown in Table 7 below.

<u>Table 7</u>		
<u>Material</u>	<u>ATCC Dep. No.</u>	<u>Deposit Date</u>
DNA26843-1389	203099	August 4, 1998
DNA30867-1335	209807	April 28, 1998
DNA34431-1177	209399	October 17, 1997

<u>Table 7</u>		
<u>Material</u>	<u>ATCC Dep. No.</u>	<u>Deposit Date</u>
DNA38268-1188	209421	October 28, 1997
DNA40621-1440	209922	June 2, 1998
DNA40625-1189	209788	April 21, 1998
DNA45409-2511	203579	January 12, 1999
DNA45495-1550	203156	August 25, 1998
DNA49820-1427	209932	June 2, 1998
DNA56406-1704	203478	November 17, 1998
DNA56410-1414	209923	June 2, 1998
DNA56436-1448	209902	May 27, 1998
DNA56855-1447	203004	June 23, 1998
DNA56860-1510	209952	June 9, 1998
DNA56862-1343	203174	September 1, 1998
DNA56868-1478	203024	June 23, 1998
DNA56869-1545	203161	August 25, 1998
DNA57704-1452	209953	June 9, 1998
DNA58723-1588	203133	August 18, 1998
DNA57827-1493	203045	July 1, 1998
DNA58737-1473	203136	August 18, 1998
DNA58846-1409	209957	June 9, 1998
DNA58850-1495	209956	June 9, 1998
DNA58855-1422	203018	June 23, 1998
DNA59211-1450	209960	June 9, 1998
DNA59212-1627	203245	September 9, 1998
DNA59213-1487	209959	June 9, 1998
DNA59605-1418	203005	June 23, 1998
DNA59609-1470	209963	June 9, 1998
DNA59610-1556	209990	June 16, 1998
DNA59837-2545	203658	February 9, 1999
DNA59844-2542	203650	February 9, 1999
DNA59854-1459	209974	June 16, 1998
DNA60625-1507	209975	June 16, 1998
DNA60629-1481	209979	June 16, 1998
DNA61755-1554	203112	August 11, 1998
DNA62812-1594	203248	September 9, 1998
DNA62815-1576	203247	September 9, 1998
DNA64881-1602	203240	September 9, 1998
DNA64886-1601	203241	September 9, 1998
DNA64902-1667	203317	October 6, 1998
DNA64950-1590	203224	September 15, 1998
DNA65403-1565	203230	September 15, 1998

<u>Table 7</u>		
<u>Material</u>	<u>ATCC Dep. No.</u>	<u>Deposit Date</u>
DNA66308-1537	203159	August 25, 1998
DNA66519-1535	203236	September 15, 1998
DNA66521-1583	203225	September 15, 1998
DNA66658-1584	203229	September 15, 1998
DNA66660-1585	203279	September 22, 1998
DNA66663-1598	203268	September 22, 1998
DNA66674-1599	203281	September 22, 1998
DNA68862-2546	203652	February 9, 1999
DNA68866-1644	203283	September 22, 1998
DNA68871-1638	203280	September 22, 1998
DNA68880-1676	203319	October 6, 1998
DNA68883-1691	203535	December 15, 1998
DNA68885-1678	203311	October 6, 1998
DNA71277-1636	203285	September 22, 1998
DNA73727-1673	203459	November 3, 1998
DNA73734-1680	203363	October 20, 1998
DNA73735-1681	203356	October 20, 1998
DNA76393-1664	203323	October 6, 1998
DNA77301-1708	203407	October 27, 1998
DNA77568-1626	203134	August 18, 1998
DNA77626-1705	203536	December 15, 1998
DNA81754-2532	203542	December 15, 1998
DNA81757-2512	203543	December 15, 1998
DNA82302-2529	203534	December 15, 1998
DNA82340-2530	203547	December 22, 1998
DNA83500-2506	203391	October 29, 1998
DNA84920-2614	203966	April 27, 1999
DNA85066-2534	203588	January 12, 1999
DNA86571-2551	203660	February 9, 1999
DNA87991-2540	203656	February 9, 1999
DNA92238-2539	203602	January 20, 1999
DNA96042-2682	PTA-382	July 20, 1999
DNA96787-2534	203589	January 12, 1999
DNA125185-2806	PTA-1031	December 7, 1999
DNA147531-2821	PTA-1185	January 11, 2000
DNA115291-2681	PTA-202	June 8, 1999
DNA164625-28890	PTA-1535	March 21, 2000
DNA131639-2874	PTA-1784	April 25, 2000
DNA79230-2525	203549	December 22, 1998



[0401] These deposits were made under the provisions of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purpose of Patent Procedure and the Regulations thereunder (Budapest Treaty). This assures maintenance of a viable culture of the deposit for 30 years from the date of deposit. The deposits will be made available by ATCC under the terms of the Budapest Treaty, and subject to an agreement between Genentech, Inc. and ATCC, which assures permanent and unrestricted availability of the progeny of the culture of the deposit to the public upon issuance of the pertinent U.S. patent or upon laying open to the public of any U.S. or foreign patent application, whichever comes first, and assures availability of the progeny to one determined by the U.S. Commissioner of Patents and Trademarks to be entitled thereto according to 35 USC § 122 and the Commissioner's rules pursuant thereto (including 37 CFR § 1.14 with particular reference to 886 OG 638).

[0402] The assignee of the present application has agreed that if a culture of the materials on deposit should die or be lost or destroyed when cultivated under suitable conditions, the materials will be promptly replaced on notification with another of the same. Availability of the deposited material is not to be construed as a license to practice the invention in contravention of the rights granted under the authority of any government in accordance with its patent laws.

#### EXAMPLE 5

##### Use of PRO as a hybridization probe

[0403] The following method describes use of a nucleotide sequence encoding PRO as a hybridization probe.

[0404] DNA comprising the coding sequence of full-length or mature PRO as disclosed herein is employed as a probe to screen for homologous DNAs (such as those encoding naturally-occurring variants of PRO) in human tissue cDNA libraries or human tissue genomic libraries.

[0405] Hybridization and washing of filters containing either library DNAs is performed under the following high stringency conditions. Hybridization of radiolabeled PRO-derived probe to the filters is performed in a solution of 50% formamide, 5x SSC, 0.1%

SDS, 0.1% sodium pyrophosphate, 50 mM sodium phosphate, pH 6.8, 2x Denhardt's solution, and 10% dextran sulfate at 42°C for 20 hours. Washing of the filters is performed in an aqueous solution of 0.1x SSC and 0.1% SDS at 42°C.

[0406] DNAs having a desired sequence identity with the DNA encoding full-length native sequence PRO can then be identified using standard techniques known in the art.

### EXAMPLE 6

#### Expression of PRO in *E. coli*

[0407] This example illustrates preparation of an unglycosylated form of PRO by recombinant expression in *E. coli*.

[0408] The DNA sequence encoding PRO is initially amplified using selected PCR primers. The primers should contain restriction enzyme sites which correspond to the restriction enzyme sites on the selected expression vector. A variety of expression vectors may be employed. An example of a suitable vector is pBR322 (derived from *E. coli*; see Bolivar et al., Gene, 2:95 (1977)) which contains genes for ampicillin and tetracycline resistance. The vector is digested with restriction enzyme and dephosphorylated. The PCR amplified sequences are then ligated into the vector. The vector will preferably include sequences which encode for an antibiotic resistance gene, a trp promoter, a polyhis leader (including the first six STII codons, polyhis sequence, and enterokinase cleavage site), the PRO coding region, lambda transcriptional terminator, and an argU gene.

[0409] The ligation mixture is then used to transform a selected *E. coli* strain using the methods described in Sambrook et al., supra. Transformants are identified by their ability to grow on LB plates and antibiotic resistant colonies are then selected. Plasmid DNA can be isolated and confirmed by restriction analysis and DNA sequencing.

[0410] Selected clones can be grown overnight in liquid culture medium such as LB broth supplemented with antibiotics. The overnight culture may subsequently be used to inoculate a larger scale culture. The cells are then grown to a desired optical density, during which the expression promoter is turned on.

[0411] After culturing the cells for several more hours, the cells can be harvested by centrifugation. The cell pellet obtained by the centrifugation can be solubilized using various agents known in the art, and the solubilized PRO protein can then be purified using a metal chelating column under conditions that allow tight binding of the protein.

[0412] PRO may be expressed in *E. coli* in a poly-His tagged form, using the following procedure. The DNA encoding PRO is initially amplified using selected PCR primers. The primers will contain restriction enzyme sites which correspond to the restriction enzyme sites on the selected expression vector, and other useful sequences providing for efficient and reliable translation initiation, rapid purification on a metal chelation column, and proteolytic removal with enterokinase. The PCR-amplified, poly-His tagged sequences are then ligated into an expression vector, which is used to transform an *E. coli* host based on strain 52 (W3110 fuhA(tonA) lon galE rpoHts(htpRts) clpP(lacIq). Transformants are first grown in LB containing 50 mg/ml carbenicillin at 30°C with shaking until an O.D.600 of 3-5 is reached. Cultures are then diluted 50-100 fold into CRAP media (prepared by mixing 3.57 g (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 0.71 g sodium citrate•2H<sub>2</sub>O, 1.07 g KCl, 5.36 g Difco yeast extract, 5.36 g Sheffield hycase SF in 500 mL water, as well as 110 mM MPOS, pH 7.3, 0.55% (w/v) glucose and 7 mM MgSO<sub>4</sub>) and grown for approximately 20-30 hours at 30°C with shaking. Samples are removed to verify expression by SDS-PAGE analysis, and the bulk culture is centrifuged to pellet the cells. Cell pellets are frozen until purification and refolding.

[0413] *E. coli* paste from 0.5 to 1 L fermentations (6-10 g pellets) is resuspended in 10 volumes (w/v) in 7 M guanidine, 20 mM Tris, pH 8 buffer. Solid sodium sulfite and sodium tetrathionate is added to make final concentrations of 0.1M and 0.02 M, respectively, and the solution is stirred overnight at 4°C. This step results in a denatured protein with all cysteine residues blocked by sulfitolization. The solution is centrifuged at 40,000 rpm in a Beckman Ultracentrifuge for 30 min. The supernatant is diluted with 3-5 volumes of metal chelate column buffer (6 M guanidine, 20 mM Tris, pH 7.4) and filtered through 0.22 micron filters to clarify. The clarified extract is loaded onto a 5 ml Qiagen Ni-NTA metal chelate column equilibrated in the metal chelate column buffer. The column is washed with additional buffer containing 50 mM imidazole (Calbiochem, Utrol grade), pH 7.4. The protein is eluted with buffer containing 250 mM imidazole. Fractions containing the desired protein are pooled and

stored at 4°C. Protein concentration is estimated by its absorbance at 280 nm using the calculated extinction coefficient based on its amino acid sequence.

**[0414]** The proteins are refolded by diluting the sample slowly into freshly prepared refolding buffer consisting of: 20 mM Tris, pH 8.6, 0.3 M NaCl, 2.5 M urea, 5 mM cysteine, 20 mM glycine and 1 mM EDTA. Refolding volumes are chosen so that the final protein concentration is between 50 to 100 micrograms/ml. The refolding solution is stirred gently at 4°C for 12-36 hours. The refolding reaction is quenched by the addition of TFA to a final concentration of 0.4% (pH of approximately 3). Before further purification of the protein, the solution is filtered through a 0.22 micron filter and acetonitrile is added to 2-10% final concentration. The refolded protein is chromatographed on a Poros R1/H reversed phase column using a mobile buffer of 0.1% TFA with elution with a gradient of acetonitrile from 10 to 80%. Aliquots of fractions with A280 absorbance are analyzed on SDS polyacrylamide gels and fractions containing homogeneous refolded protein are pooled. Generally, the properly refolded species of most proteins are eluted at the lowest concentrations of acetonitrile since those species are the most compact with their hydrophobic interiors shielded from interaction with the reversed phase resin. Aggregated species are usually eluted at higher acetonitrile concentrations. In addition to resolving misfolded forms of proteins from the desired form, the reversed phase step also removes endotoxin from the samples.

**[0415]** Fractions containing the desired folded PRO polypeptide are pooled and the acetonitrile removed using a gentle stream of nitrogen directed at the solution. Proteins are formulated into 20 mM Hepes, pH 6.8 with 0.14 M sodium chloride and 4% mannitol by dialysis or by gel filtration using G25 Superfine (Pharmacia) resins equilibrated in the formulation buffer and sterile filtered.

**[0416]** Many of the PRO polypeptides disclosed herein were successfully expressed as described above.

## EXAMPLE 7

### Expression of PRO in mammalian cells

[0417] This example illustrates preparation of a potentially glycosylated form of PRO by recombinant expression in mammalian cells.

[0418] The vector, pRK5 (see EP 307,247, published March 15, 1989), is employed as the expression vector. Optionally, the PRO DNA is ligated into pRK5 with selected restriction enzymes to allow insertion of the PRO DNA using ligation methods such as described in Sambrook et al., supra. The resulting vector is called pRK5-PRO.

[0419] In one embodiment, the selected host cells may be 293 cells. Human 293 cells (ATCC CCL 1573) are grown to confluence in tissue culture plates in medium such as DMEM supplemented with fetal calf serum and optionally, nutrient components and/or antibiotics. About 10 µg pRK5-PRO DNA is mixed with about 1 µg DNA encoding the VA RNA gene [Thimmappaya et al., Cell, 31:543 (1982)] and dissolved in 500 µl of 1 mM Tris-HCl, 0.1 mM EDTA, 0.227 M CaCl<sub>2</sub>. To this mixture is added, dropwise, 500 µl of 50 mM HEPES (pH 7.35), 280 mM NaCl, 1.5 mM NaPO<sub>4</sub>, and a precipitate is allowed to form for 10 minutes at 25°C. The precipitate is suspended and added to the 293 cells and allowed to settle for about four hours at 37°C. The culture medium is aspirated off and 2 ml of 20% glycerol in PBS is added for 30 seconds. The 293 cells are then washed with serum free medium, fresh medium is added and the cells are incubated for about 5 days.

[0420] Approximately 24 hours after the transfections, the culture medium is removed and replaced with culture medium (alone) or culture medium containing 200 µCi/ml <sup>35</sup>S-cysteine and 200 µCi/ml <sup>35</sup>S-methionine. After a 12 hour incubation, the conditioned medium is collected, concentrated on a spin filter, and loaded onto a 15% SDS gel. The processed gel may be dried and exposed to film for a selected period of time to reveal the presence of PRO polypeptide. The cultures containing transfected cells may undergo further incubation (in serum free medium) and the medium is tested in selected bioassays.

[0421] In an alternative technique, PRO may be introduced into 293 cells transiently using the dextran sulfate method described by Somparyrac et al., Proc. Natl. Acad. Sci., 12:7575 (1981). 293 cells are grown to maximal density in a spinner flask and 700 µg pRK5-PRO DNA is added. The cells are first concentrated from the spinner flask by

centrifugation and washed with PBS. The DNA-dextran precipitate is incubated on the cell pellet for four hours. The cells are treated with 20% glycerol for 90 seconds, washed with tissue culture medium, and re-introduced into the spinner flask containing tissue culture medium, 5 µg/ml bovine insulin and 0.1 µg/ml bovine transferrin. After about four days, the conditioned media is centrifuged and filtered to remove cells and debris. The sample containing expressed PRO can then be concentrated and purified by any selected method, such as dialysis and/or column chromatography.

**[0422]** In another embodiment, PRO can be expressed in CHO cells. The pRK5-PRO can be transfected into CHO cells using known reagents such as CaPO<sub>4</sub> or DEAE-dextran. As described above, the cell cultures can be incubated, and the medium replaced with culture medium (alone) or medium containing a radiolabel such as <sup>35</sup>S-methionine. After determining the presence of PRO polypeptide, the culture medium may be replaced with serum free medium. Preferably, the cultures are incubated for about 6 days, and then the conditioned medium is harvested. The medium containing the expressed PRO can then be concentrated and purified by any selected method.

**[0423]** Epitope-tagged PRO may also be expressed in host CHO cells. The PRO may be subcloned out of the pRK5 vector. The subclone insert can undergo PCR to fuse in frame with a selected epitope tag such as a poly-his tag into a Baculovirus expression vector. The poly-his tagged PRO insert can then be subcloned into a SV40 driven vector containing a selection marker such as DHFR for selection of stable clones. Finally, the CHO cells can be transfected (as described above) with the SV40 driven vector. Labeling may be performed, as described above, to verify expression. The culture medium containing the expressed poly-His tagged PRO can then be concentrated and purified by any selected method, such as by Ni<sup>2+</sup>-chelate affinity chromatography.

**[0424]** PRO may also be expressed in CHO and/or COS cells by a transient expression procedure or in CHO cells by another stable expression procedure.

**[0425]** Stable expression in CHO cells is performed using the following procedure. The proteins are expressed as an IgG construct (immunoadhesin), in which the coding sequences for the soluble forms (e.g. extracellular domains) of the respective proteins are fused

to an IgG1 constant region sequence containing the hinge, CH2 and CH2 domains and/or is a poly-His tagged form.

**[0426]** Following PCR amplification, the respective DNAs are subcloned in a CHO expression vector using standard techniques as described in Ausubel et al., Current Protocols of Molecular Biology, Unit 3.16, John Wiley and Sons (1997). CHO expression vectors are constructed to have compatible restriction sites 5' and 3' of the DNA of interest to allow the convenient shuttling of cDNA's. The vector used expression in CHO cells is as described in Lucas et al., Nucl. Acids Res. 24:9 (1774-1779 (1996), and uses the SV40 early promoter/enhancer to drive expression of the cDNA of interest and dihydrofolate reductase (DHFR). DHFR expression permits selection for stable maintenance of the plasmid following transfection.

**[0427]** Twelve micrograms of the desired plasmid DNA is introduced into approximately 10 million CHO cells using commercially available transfection reagents Superfect<sup>®</sup> SUPERFECT<sup>™</sup> (Quiagen), Desper<sup>®</sup> DOSPER<sup>™</sup> or Fugene<sup>®</sup> FUGENE<sup>™</sup> (Boehringer Mannheim). The cells are grown as described in Lucas et al., supra. Approximately  $3 \times 10^{-7}$  cells are frozen in an ampule for further growth and production as described below.

**[0428]** The ampules containing the plasmid DNA are thawed by placement into water bath and mixed by vortexing. The contents are pipetted into a centrifuge tube containing 10 mLs of media and centrifuged at 1000 rpm for 5 minutes. The supernatant is aspirated and the cells are resuspended in 10 mL of selective media (0.2  $\mu$ m filtered PS20 with 5% 0.2  $\mu$ m diafiltered fetal bovine serum). The cells are then aliquoted into a 100 mL spinner containing 90 mL of selective media. After 1-2 days, the cells are transferred into a 250 mL spinner filled with 150 mL selective growth medium and incubated at 37°C. After another 2-3 days, 250 mL, 500 mL and 2000 mL spinners are seeded with  $3 \times 10^5$  cells/mL. The cell media is exchanged with fresh media by centrifugation and resuspension in production medium. Although any suitable CHO media may be employed, a production medium described in U.S. Patent No. 5,122,469, issued June 16, 1992 may actually be used. A 3L production spinner is seeded at  $1.2 \times 10^6$  cells/mL. On day 0, the cell number pH is determined. On day 1, the spinner is sampled and sparging with filtered air is commenced. On day 2, the spinner is sampled, the

temperature shifted to 33°C, and 30 mL of 500 g/L glucose and 0.6 mL of 10% antifoam (e.g., 35% polydimethylsiloxane emulsion, Dow Corning 365 Medical Grade Emulsion) taken. Throughout the production, the pH is adjusted as necessary to keep it at around 7.2. After 10 days, or until the viability dropped below 70%, the cell culture is harvested by centrifugation and filtering through a 0.22 µm filter. The filtrate was either stored at 4°C or immediately loaded onto columns for purification.

[0429] For the poly-His tagged constructs, the proteins are purified using a Ni-NTA column (Qiagen). Before purification, imidazole is added to the conditioned media to a concentration of 5 mM. The conditioned media is pumped onto a 6 ml Ni-NTA column equilibrated in 20 mM Hepes, pH 7.4, buffer containing 0.3 M NaCl and 5 mM imidazole at a flow rate of 4-5 ml/min. at 4°C. After loading, the column is washed with additional equilibration buffer and the protein eluted with equilibration buffer containing 0.25 M imidazole. The highly purified protein is subsequently desalted into a storage buffer containing 10 mM Hepes, 0.14 M NaCl and 4% mannitol, pH 6.8, with a 25 ml G25 Superfine (Pharmacia) column and stored at -80°C.

[0430] Immunoadhesin (Fc-containing) constructs are purified from the conditioned media as follows. The conditioned medium is pumped onto a 5 ml Protein A column (Pharmacia) which had been equilibrated in 20 mM Na phosphate buffer, pH 6.8. After loading, the column is washed extensively with equilibration buffer before elution with 100 mM citric acid, pH 3.5. The eluted protein is immediately neutralized by collecting 1 ml fractions into tubes containing 275 µL of 1 M Tris buffer, pH 9. The highly purified protein is subsequently desalted into storage buffer as described above for the poly-His tagged proteins. The homogeneity is assessed by SDS polyacrylamide gels and by N-terminal amino acid sequencing by Edman degradation.

[0431] Many of the PRO polypeptides disclosed herein were successfully expressed as described above.

## EXAMPLE 8

### Expression of PRO in Yeast

[0432] The following method describes recombinant expression of PRO in yeast.



[0433] First, yeast expression vectors are constructed for intracellular production or secretion of PRO from the ADH2/GAPDH promoter. DNA encoding PRO and the promoter is inserted into suitable restriction enzyme sites in the selected plasmid to direct intracellular expression of PRO. For secretion, DNA encoding PRO can be cloned into the selected plasmid, together with DNA encoding the ADH2/GAPDH promoter, a native PRO signal peptide or other mammalian signal peptide, or, for example, a yeast alpha-factor or invertase secretory signal/leader sequence, and linker sequences (if needed) for expression of PRO.

[0434] Yeast cells, such as yeast strain AB110, can then be transformed with the expression plasmids described above and cultured in selected fermentation media. The transformed yeast supernatants can be analyzed by precipitation with 10% trichloroacetic acid and separation by SDS-PAGE, followed by staining of the gels with Coomassie Blue stain.

[0435] Recombinant PRO can subsequently be isolated and purified by removing the yeast cells from the fermentation medium by centrifugation and then concentrating the medium using selected cartridge filters. The concentrate containing PRO may further be purified using selected column chromatography resins.

[0436] Many of the PRO polypeptides disclosed herein were successfully expressed as described above.

## EXAMPLE 9

### Expression of PRO in Baculovirus-Infected Insect Cells

[0437] The following method describes recombinant expression of PRO in Baculovirus-infected insect cells.

[0438] The sequence coding for PRO is fused upstream of an epitope tag contained within a baculovirus expression vector. Such epitope tags include poly-his tags and immunoglobulin tags (like Fc regions of IgG). A variety of plasmids may be employed, including plasmids derived from commercially available plasmids such as pVL1393 (Novagen). Briefly, the sequence encoding PRO or the desired portion of the coding sequence of PRO such as the sequence encoding the extracellular domain of a transmembrane protein or the sequence encoding the mature protein if the protein is extracellular is amplified by PCR

with primers complementary to the 5' and 3' regions. The 5' primer may incorporate flanking (selected) restriction enzyme sites. The product is then digested with those selected restriction enzymes and subcloned into the expression vector.

[0439] Recombinant baculovirus is generated by co-transfecting the above plasmid and ~~BaculoGold™~~ BACULOGOLD™ virus DNA (Pharmingen) into *Spodoptera frugiperda* ("Sf9") cells (ATCC CRL 1711) using ~~lipofectin~~ LIPOFECTIN cationic lipid (commercially available from GIBCO-BRL). After 4 - 5 days of incubation at 28°C, the released viruses are harvested and used for further amplifications. Viral infection and protein expression are performed as described by O'Reilley et al., Baculovirus expression vectors: A Laboratory Manual, Oxford: Oxford University Press (1994).

[0440] Expressed poly-his tagged PRO can then be purified, for example, by Ni<sup>2+</sup>-chelate affinity chromatography as follows. Extracts are prepared from recombinant virus-infected Sf9 cells as described by Rupert et al., Nature, 362:175-179 (1993). Briefly, Sf9 cells are washed, resuspended in sonication buffer (25 mL Hepes, pH 7.9; 12.5 mM MgCl<sub>2</sub>; 0.1 mM EDTA; 10% glycerol; 0.1% NP-40; 0.4 M KCl), and sonicated twice for 20 seconds on ice. The sonicates are cleared by centrifugation, and the supernatant is diluted 50-fold in loading buffer (50 mM phosphate, 300 mM NaCl, 10% glycerol, pH 7.8) and filtered through a 0.45 µm filter. A Ni<sup>2+</sup>-NTA agarose column (commercially available from Qiagen) is prepared with a bed volume of 5 mL, washed with 25 mL of water and equilibrated with 25 mL of loading buffer. The filtered cell extract is loaded onto the column at 0.5 mL per minute. The column is washed to baseline A<sub>280</sub> with loading buffer, at which point fraction collection is started. Next, the column is washed with a secondary wash buffer (50 mM phosphate; 300 mM NaCl, 10% glycerol, pH 6.0), which elutes nonspecifically bound protein. After reaching A<sub>280</sub> baseline again, the column is developed with a 0 to 500 mM Imidazole gradient in the secondary wash buffer. One mL fractions are collected and analyzed by SDS-PAGE and silver staining or Western blot with Ni<sup>2+</sup>-NTA-conjugated to alkaline phosphatase (Qiagen). Fractions containing the eluted His<sub>10</sub>-tagged PRO are pooled and dialyzed against loading buffer.

[0441] Alternatively, purification of the IgG tagged (or Fc tagged) PRO can be performed using known chromatography techniques, including for instance, Protein A or protein G column chromatography.

[0442] Many of the PRO polypeptides disclosed herein were successfully expressed as described above.

#### EXAMPLE 10

##### Preparation of Antibodies that Bind PRO

[0443] This example illustrates preparation of monoclonal antibodies which can specifically bind PRO.

[0444] Techniques for producing the monoclonal antibodies are known in the art and are described, for instance, in Goding, supra. Immunogens that may be employed include purified PRO, fusion proteins containing PRO, and cells expressing recombinant PRO on the cell surface. Selection of the immunogen can be made by the skilled artisan without undue experimentation.

[0445] Mice, such as Balb/c, are immunized with the PRO immunogen emulsified in complete Freund's adjuvant and injected subcutaneously or intraperitoneally in an amount from 1-100 micrograms. Alternatively, the immunogen is emulsified in MPL-TDM adjuvant (Ribi Immunochemical Research, Hamilton, MT) and injected into the animal's hind foot pads. The immunized mice are then boosted 10 to 12 days later with additional immunogen emulsified in the selected adjuvant. Thereafter, for several weeks, the mice may also be boosted with additional immunization injections. Serum samples may be periodically obtained from the mice by retro-orbital bleeding for testing in ELISA assays to detect anti-PRO antibodies.

[0446] After a suitable antibody titer has been detected, the animals "positive" for antibodies can be injected with a final intravenous injection of PRO. Three to four days later, the mice are sacrificed and the spleen cells are harvested. The spleen cells are then fused (using 35% polyethylene glycol) to a selected murine myeloma cell line such as P3X63AgU.1, available from ATCC, No. CRL 1597. The fusions generate hybridoma cells which can then be plated in 96 well tissue culture plates containing HAT (hypoxanthine, aminopterin, and

thymidine) medium to inhibit proliferation of non-fused cells, myeloma hybrids, and spleen cell hybrids.

[0447] The hybridoma cells will be screened in an ELISA for reactivity against PRO. Determination of “positive” hybridoma cells secreting the desired monoclonal antibodies against PRO is within the skill in the art.

[0448] The positive hybridoma cells can be injected intraperitoneally into syngeneic Balb/c mice to produce ascites containing the anti-PRO monoclonal antibodies. Alternatively, the hybridoma cells can be grown in tissue culture flasks or roller bottles. Purification of the monoclonal antibodies produced in the ascites can be accomplished using ammonium sulfate precipitation, followed by gel exclusion chromatography. Alternatively, affinity chromatography based upon binding of antibody to protein A or protein G can be employed.

#### EXAMPLE 11: Purification of PRO Polypeptides Using Specific Antibodies

[0449] Native or recombinant PRO polypeptides may be purified by a variety of standard techniques in the art of protein purification. For example, pro-PRO polypeptide, mature PRO polypeptide, or pre-PRO polypeptide is purified by immunoaffinity chromatography using antibodies specific for the PRO polypeptide of interest. In general, an immunoaffinity column is constructed by covalently coupling the anti-PRO polypeptide antibody to an activated chromatographic resin.

[0450] Polyclonal immunoglobulins are prepared from immune sera either by precipitation with ammonium sulfate or by purification on immobilized Protein A (Pharmacia LKB Biotechnology, Piscataway, N.J.). Likewise, monoclonal antibodies are prepared from mouse ascites fluid by ammonium sulfate precipitation or chromatography on immobilized Protein A. Partially purified immunoglobulin is covalently attached to a chromatographic resin such as CnBr-activated SEPHAROSE™ (Pharmacia LKB Biotechnology). The antibody is coupled to the resin, the resin is blocked, and the derivative resin is washed according to the manufacturer's instructions.

[0451] Such an immunoaffinity column is utilized in the purification of PRO polypeptide by preparing a fraction from cells containing PRO polypeptide in a soluble form.

This preparation is derived by solubilization of the whole cell or of a subcellular fraction obtained via differential centrifugation by the addition of detergent or by other methods well known in the art. Alternatively, soluble PRO polypeptide containing a signal sequence may be secreted in useful quantity into the medium in which the cells are grown.

[0452] A soluble PRO polypeptide-containing preparation is passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of PRO polypeptide (*e.g.*, high ionic strength buffers in the presence of detergent). Then, the column is eluted under conditions that disrupt antibody/PRO polypeptide binding (*e.g.*, a low pH buffer such as approximately pH 2-3, or a high concentration of a chaotrope such as urea or thiocyanate ion), and PRO polypeptide is collected.

#### EXAMPLE 12: Drug Screening

[0453] This invention is particularly useful for screening compounds by using PRO polypeptides or binding fragment thereof in any of a variety of drug screening techniques. The PRO polypeptide or fragment employed in such a test may either be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. One method of drug screening utilizes eukaryotic or prokaryotic host cells which are stably transformed with recombinant nucleic acids expressing the PRO polypeptide or fragment. Drugs are screened against such transformed cells in competitive binding assays. Such cells, either in viable or fixed form, can be used for standard binding assays. One may measure, for example, the formation of complexes between PRO polypeptide or a fragment and the agent being tested. Alternatively, one can examine the diminution in complex formation between the PRO polypeptide and its target cell or target receptors caused by the agent being tested.

[0454] Thus, the present invention provides methods of screening for drugs or any other agents which can affect a PRO polypeptide-associated disease or disorder. These methods comprise contacting such an agent with an PRO polypeptide or fragment thereof and assaying (i) for the presence of a complex between the agent and the PRO polypeptide or fragment, or (ii) for the presence of a complex between the PRO polypeptide or fragment and the cell, by methods well known in the art. In such competitive binding assays, the PRO polypeptide or fragment is typically labeled. After suitable incubation, free PRO polypeptide

or fragment is separated from that present in bound form, and the amount of free or uncomplexed label is a measure of the ability of the particular agent to bind to PRO polypeptide or to interfere with the PRO polypeptide/cell complex.

[0455] Another technique for drug screening provides high throughput screening for compounds having suitable binding affinity to a polypeptide and is described in detail in WO 84/03564, published on September 13, 1984. Briefly stated, large numbers of different small peptide test compounds are synthesized on a solid substrate, such as plastic pins or some other surface. As applied to a PRO polypeptide, the peptide test compounds are reacted with PRO polypeptide and washed. Bound PRO polypeptide is detected by methods well known in the art. Purified PRO polypeptide can also be coated directly onto plates for use in the aforementioned drug screening techniques. In addition, non-neutralizing antibodies can be used to capture the peptide and immobilize it on the solid support.

[0456] This invention also contemplates the use of competitive drug screening assays in which neutralizing antibodies capable of binding PRO polypeptide specifically compete with a test compound for binding to PRO polypeptide or fragments thereof. In this manner, the antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with PRO polypeptide.

### EXAMPLE 13

#### Rational Drug Design

[0457] The goal of rational drug design is to produce structural analogs of biologically active polypeptide of interest (*i.e.*, a PRO polypeptide) or of small molecules with which they interact, *e.g.*, agonists, antagonists, or inhibitors. Any of these examples can be used to fashion drugs which are more active or stable forms of the PRO polypeptide or which enhance or interfere with the function of the PRO polypeptide *in vivo* (*c.f.*, Hodgson, Bio/Technology, 9: 19-21 (1991)).

[0458] In one approach, the three-dimensional structure of the PRO polypeptide, or of an PRO polypeptide-inhibitor complex, is determined by x-ray crystallography, by computer modeling or, most typically, by a combination of the two approaches. Both the shape and charges of the PRO polypeptide must be ascertained to elucidate the structure and to determine

active site(s) of the molecule. Less often, useful information regarding the structure of the PRO polypeptide may be gained by modeling based on the structure of homologous proteins. In both cases, relevant structural information is used to design analogous PRO polypeptide-like molecules or to identify efficient inhibitors. Useful examples of rational drug design may include molecules which have improved activity or stability as shown by Braxton and Wells, Biochemistry, 31:7796-7801 (1992) or which act as inhibitors, agonists, or antagonists of native peptides as shown by Athauda *et al.*, J. Biochem., 113:742-746 (1993).

[0459] It is also possible to isolate a target-specific antibody, selected by functional assay, as described above, and then to solve its crystal structure. This approach, in principle, yields a pharmacore upon which subsequent drug design can be based. It is possible to bypass protein crystallography altogether by generating anti-idiotypic antibodies (anti-ids) to a functional, pharmacologically active antibody. As a mirror image of a mirror image, the binding site of the anti-ids would be expected to be an analog of the original receptor. The anti-id could then be used to identify and isolate peptides from banks of chemically or biologically produced peptides. The isolated peptides would then act as the pharmacore.

[0460] By virtue of the present invention, sufficient amounts of the PRO polypeptide may be made available to perform such analytical studies as X-ray crystallography. In addition, knowledge of the PRO polypeptide amino acid sequence provided herein will provide guidance to those employing computer modeling techniques in place of or in addition to x-ray crystallography.

#### EXAMPLE 14

##### Pericyte c-Fos Induction (Assay 93)

[0461] This assay shows that certain polypeptides of the invention act to induce the expression of c-fos in pericyte cells and, therefore, are useful not only as diagnostic markers for particular types of pericyte-associated tumors but also for giving rise to antagonists which would be expected to be useful for the therapeutic treatment of pericyte-associated tumors. Induction of c-fos expression in pericytes is also indicative of the induction of angiogenesis and, as such, PRO polypeptides capable of inducing the expression of c-fos would be expected to be useful for the treatment of conditions where induced angiogenesis would be beneficial

including, for example, wound healing, and the like. Specifically, on day 1, pericytes are received from VEC Technologies and all but 5 ml of media is removed from flask. On day 2, the pericytes are trypsinized, washed, spun and then plated onto 96 well plates. On day 7, the media is removed and the pericytes are treated with 100 µl of PRO polypeptide test samples and controls (positive control = DME+5% serum +/- PDGF at 500 ng/ml; negative control = protein 32). Replicates are averaged and SD/CV are determined. Fold increase over Protein 32 (buffer control) value indicated by chemiluminescence units (RLU) luminometer reading verses frequency is plotted on a histogram. Two-fold above Protein 32 value is considered positive for the assay. ASY Matrix: Growth media = low glucose DMEM = 20% FBS + 1X pen strep + 1X fungizone. Assay Media = low glucose DMEM +5% FBS.

[0462] The following polypeptides tested positive in this assay: PRO1347 and PRO1340.

#### EXAMPLE 15

##### Ability of PRO Polypeptides to Stimulate the Release of Proteoglycans from Cartilage

##### (Assay 97)

[0463] The ability of various PRO polypeptides to stimulate the release of proteoglycans from cartilage tissue was tested as follows.

[0464] The metacarpophalangeal joint of 4-6 month old pigs was aseptically dissected, and articular cartilage was removed by free hand slicing being careful to avoid the underlying bone. The cartilage was minced and cultured in bulk for 24 hours in a humidified atmosphere of 95% air, 5% CO<sub>2</sub> in serum free (SF) media (DME/F12 1:1) with 0.1% BSA and 100U/ml penicillin and 100µg/ml streptomycin. After washing three times, approximately 100 mg of articular cartilage was aliquoted into micronics tubes and incubated for an additional 24 hours in the above SF media. PRO polypeptides were then added at 1% either alone or in combination with 18 ng/ml interleukin-1a, a known stimulator of proteoglycan release from cartilage tissue. The supernatant was then harvested and assayed for the amount of proteoglycans using the 1,9-dimethyl-methylene blue (DMB) colorimetric assay (Farndale and Buttle, Biochem. Biophys. Acta 883:173-177 (1985)). A positive result in this assay indicates



that the test polypeptide will find use, for example, in the treatment of sports-related joint problems, articular cartilage defects, osteoarthritis or rheumatoid arthritis.

[0465] When various PRO polypeptides were tested in the above assay, the polypeptides demonstrated a marked ability to stimulate release of proteoglycans from cartilage tissue both basally and after stimulation with interleukin-1 $\alpha$  and at 24 and 72 hours after treatment, thereby indicating that these PRO polypeptides are useful for stimulating proteoglycan release from cartilage tissue. As such, these PRO polypeptides are useful for the treatment of sports-related joint problems, articular cartilage defects, osteoarthritis or rheumatoid arthritis. The polypeptides testing positive in this assay are: PRO1565, PRO1693, PRO1801 and PRO10096.

#### EXAMPLE 16

##### Detection of Polypeptides That Affect Glucose or FFA Uptake in Skeletal Muscle (Assay 106)

[0466] This assay is designed to determine whether PRO polypeptides show the ability to affect glucose or FFA uptake by skeletal muscle cells. PRO polypeptides testing positive in this assay would be expected to be useful for the therapeutic treatment of disorders where either the stimulation or inhibition of glucose uptake by skeletal muscle would be beneficial including, for example, diabetes or hyper- or hypo-insulinemia.

[0467] In a 96 well format, PRO polypeptides to be assayed are added to primary rat differentiated skeletal muscle, and allowed to incubate overnight. Then fresh media with the PRO polypeptide and +/- insulin are added to the wells. The sample media is then monitored to determine glucose and FFA uptake by the skeletal muscle cells. The insulin will stimulate glucose and FFA uptake by the skeletal muscle, and insulin in media without the PRO polypeptide is used as a positive control, and a limit for scoring. As the PRO polypeptide being tested may either stimulate or inhibit glucose and FFA uptake, results are scored as positive in the assay if greater than 1.5 times or less than 0.5 times the insulin control.

[0468] The following PRO polypeptides tested positive as either stimulators or inhibitors of glucose and/or FFA uptake in this assay: PRO4405.

## EXAMPLE 17

### Identification of PRO Polypeptides That Stimulate TNF-a Release In Human Blood

#### (Assay 128)

[0469] This assay shows that certain PRO polypeptides of the present invention act to stimulate the release of TNF-a in human blood. PRO polypeptides testing positive in this assay are useful for, among other things, research purposes where stimulation of the release of TNF-a would be desired and for the therapeutic treatment of conditions wherein enhanced TNF-a release would be beneficial. Specifically, 200 µl of human blood supplemented with 50mM Hepes buffer (pH 7.2) is aliquotted per well in a 96 well test plate. To each well is then added 300µl of either the test PRO polypeptide in 50 mM Hepes buffer (at various concentrations) or 50 mM Hepes buffer alone (negative control) and the plates are incubated at 37°C for 6 hours. The samples are then centrifuged and 50µl of plasma is collected from each well and tested for the presence of TNF-a by ELISA assay. A positive in the assay is a higher amount of TNF-a in the PRO polypeptide treated samples as compared to the negative control samples.

[0470] The following PRO polypeptides tested positive in this assay: PRO263, PRO295, PRO1282, PRO1063, PRO1356, PRO3543, and PRO5990.

## EXAMPLE 18

### Tumor Versus Normal Differential Tissue Expression Distribution

[0471] Oligonucleotide probes were constructed from some of the PRO polypeptide-encoding nucleotide sequences shown in the accompanying figures for use in quantitative PCR amplification reactions. The oligonucleotide probes were chosen so as to give an approximately 200-600 base pair amplified fragment from the 3' end of its associated template in a standard PCR reaction. The oligonucleotide probes were employed in standard quantitative PCR amplification reactions with cDNA libraries isolated from different human tumor and normal human tissue samples and analyzed by agarose gel electrophoresis so as to obtain a quantitative determination of the level of expression of the PRO polypeptide-encoding nucleic acid in the various tumor and normal tissues tested. β-actin was used as a control to assure that equivalent amounts of nucleic acid was used in each reaction. Identification of the

differential expression of the PRO polypeptide-encoding nucleic acid in one or more tumor tissues as compared to one or more normal tissues of the same tissue type renders the molecule useful diagnostically for the determination of the presence or absence of tumor in a subject suspected of possessing a tumor as well as therapeutically as a target for the treatment of a tumor in a subject possessing such a tumor. These assays provided the following results.

<u>Molecule</u>	<u>is more highly expressed in:</u>	<u>as compared to:</u>
DNA26843-1389	normal lung	lung tumor
	rectum tumor	normal rectum
DNA30867-1335	normal kidney	kidney tumor
DNA40621-1440	normal lung	lung tumor
DNA40625-1189	normal lung	lung tumor
DNA45409-2511	melanoma tumor	normal skin
DNA56406-1704	kidney tumor	normal kidney
	normal skin	melanoma tumor
DNA56410-1414	normal stomach	stomach tumor
DNA56436-1448	normal skin	melanoma tumor
DNA56855-1447	normal esophagus	esophageal tumor
	rectum tumor	normal rectum
DNA56860-1510	normal kidney	kidney tumor
	rectum tumor	normal rectum
DNA56862-1343	kidney tumor	normal kidney
	normal lung	lung tumor
DNA56868-1478	normal stomach	stomach tumor
	normal lung	lung tumor
DNA56869-1545	normal esophagus	esophageal tumor
	normal skin	melanoma tumor

<u>Molecule</u>	<u>is more highly expressed in:</u>	<u>as compared to:</u>
DNA57704-1452	normal stomach	stomach tumor
	rectum tumor	normal rectum
DNA58723-1588	normal stomach	stomach tumor
	kidney tumor	normal kidney
	normal skin	melanoma tumor
DNA57827-1493	normal stomach	stomach tumor
	normal skin	melanoma tumor
DNA58737-1473	esophageal tumor	normal esophagus
	normal stomach	stomach tumor
DNA58846-1409	lung tumor	normal lung
DNA58850-1495	esophageal tumor	normal esophagus
	kidney tumor	normal kidney
DNA58855-1422	normal stomach	stomach tumor
	rectum tumor	normal rectum
DNA59211-1450	normal kidney	kidney tumor
DNA59212-1627	normal skin	melanoma tumor
DNA59213-1487	normal stomach	stomach tumor
	normal skin	melanoma tumor
DNA59605-1418	melanoma tumor	normal skin
DNA59609-1470	esophageal tumor	normal esophagus
DNA59610-1556	esophageal tumor	normal esophagus
	lung tumor	normal lung
	normal skin	melanoma tumor
DNA59837-2545	normal skin	melanoma tumor
DNA59844-2542	normal skin	melanoma tumor
	esophageal tumor	normal esophagus
DNA59854-1459	normal esophagus	esophageal tumor

<u>Molecule</u>	<u>is more highly expressed in:</u>	<u>as compared to:</u>
	stomach tumor	normal stomach
	normal lung	lung tumor
DNA60625-1507	normal lung	lung tumor
DNA60629-1481	normal esophagus	esophageal tumor
	normal rectum	rectum tumor
DNA61755-1554	normal stomach	stomach tumor
	kidney tumor	normal kidney
DNA62812-1594	normal stomach	stomach tumor
	normal lung	lung tumor
	normal rectum	rectum tumor
	normal skin	melanoma tumor
DNA62815-1576	esophageal tumor	normal esophagus
DNA64881-1602	normal stomach	stomach tumor
	normal lung	lung tumor
DNA64902-1667	esophageal tumor	normal esophagus
	kidney tumor	normal kidney
DNA65403-1565	normal esophagus	esophageal tumor
DNA66308-1537	normal lung	lung tumor
DNA66519-1535	kidney tumor	normal kidney
DNA66521-1583	normal esophagus	esophageal tumor
	normal stomach	stomach tumor
	normal lung	lung tumor
	normal rectum	rectum tumor
	normal skin	melanoma tumor
DNA66658-1584	normal lung	lung tumor
	melanoma tumor	normal skin
DNA66660-1585	lung tumor	normal lung
DNA66674-1599	kidney tumor	normal kidney
	normal lung	lung tumor

<u>Molecule</u>	<u>is more highly expressed in:</u>	<u>as compared to:</u>
DNA68862-2546	melanoma tumor	normal skin
DNA68866-1644	normal stomach	stomach tumor
DNA68871-1638	lung tumor	normal lung
	normal skin	melanoma tumor
DNA68880-1676	normal lung	lung tumor
	normal skin	melanoma tumor
DNA68883-1691	esophageal tumor	normal esophagus
DNA68885-1678	lung tumor	normal lung
DNA71277-1636	normal stomach	stomach tumor
DNA73734-1680	normal lung	lung tumor
DNA73735-1681	esophageal tumor	normal esophagus
	normal kidney	kidney tumor
	lung tumor	normal lung
	normal skin	melanoma tumor
DNA76393-1664	esophageal tumor	normal esophagus
	stomach tumor	normal stomach
	lung tumor	normal lung
	rectum tumor	normal rectum
DNA77568-1626	normal stomach	stomach tumor
	lung tumor	normal lung
DNA77626-1705	normal rectum	rectum tumor
DNA81754-2532	normal skin	melanoma tumor
DNA81757-2512	esophageal tumor	normal esophagus
	normal stomach	stomach tumor
	melanoma tumor	normal skin
DNA82302-2529	normal stomach	stomach tumor
	normal lung	lung tumor

<u>Molecule</u>	<u>is more highly expressed in:</u>	<u>as compared to:</u>
DNA82340-2530	normal esophagus	esophageal tumor
DNA85066-2534	lung tumor	normal lung
	normal skin	melanoma tumor
DNA87991-2540	esophageal tumor	normal esophagus
DNA92238-2539	normal skin	melanoma tumor
DNA96787-2534	normal kidney	kidney tumor

### EXAMPLE 19

#### Identification of Receptor/Ligand Interactions

**[0472]** In this assay, various PRO polypeptides are tested for ability to bind to a panel of potential receptor or ligand molecules for the purpose of identifying receptor/ligand interactions. The identification of a ligand for a known receptor, a receptor for a known ligand or a novel receptor/ligand pair is useful for a variety of indications including, for example, targeting bioactive molecules (linked to the ligand or receptor) to a cell known to express the receptor or ligand, use of the receptor or ligand as a reagent to detect the presence of the ligand or receptor in a composition suspected of containing the same, wherein the composition may comprise cells suspected of expressing the ligand or receptor, modulating the growth of or another biological or immunological activity of a cell known to express or respond to the receptor or ligand, modulating the immune response of cells or toward cells that express the receptor or ligand, allowing the preparation of agonists, antagonists and/or antibodies directed against the receptor or ligand which will modulate the growth of or a biological or immunological activity of a cell expressing the receptor or ligand, and various other indications which will be readily apparent to the ordinarily skilled artisan.

**[0473]** The assay is performed as follows. A PRO polypeptide of the present invention suspected of being a ligand for a receptor is expressed as a fusion protein containing the Fc domain of human IgG (an immunoadhesin). Receptor-ligand binding is detected by allowing interaction of the immunoadhesin polypeptide with cells (e.g. Cos cells) expressing candidate PRO polypeptide receptors and visualization of bound immunoadhesin with

fluorescent reagents directed toward the Fc fusion domain and examination by microscope. Cells expressing candidate receptors are produced by transient transfection, in parallel, of defined subsets of a library of cDNA expression vectors encoding PRO polypeptides that may function as receptor molecules. Cells are then incubated for 1 hour in the presence of the PRO polypeptide immunoadhesin being tested for possible receptor binding. The cells are then washed and fixed with paraformaldehyde. The cells are then incubated with fluorescent conjugated antibody directed against the Fc portion of the PRO polypeptide immunoadhesin (e.g. FITC conjugated goat anti-human-Fc antibody). The cells are then washed again and examined by microscope. A positive interaction is judged by the presence of fluorescent labeling of cells transfected with cDNA encoding a particular PRO polypeptide receptor or pool of receptors and an absence of similar fluorescent labeling of similarly prepared cells that have been transfected with other cDNA or pools of cDNA. If a defined pool of cDNA expression vectors is judged to be positive for interaction with a PRO polypeptide immunoadhesin, the individual cDNA species that comprise the pool are tested individually (the pool is “broken down”) to determine the specific cDNA that encodes a receptor able to interact with the PRO polypeptide immunoadhesin.

[0474] In another embodiment of this assay, an epitope-tagged potential ligand PRO polypeptide (e.g. 8 histidine “His” tag) is allowed to interact with a panel of potential receptor PRO polypeptide molecules that have been expressed as fusions with the Fc domain of human IgG (immunoadhesins). Following a 1 hour co-incubation with the epitope tagged PRO polypeptide, the candidate receptors are each immunoprecipitated with protein A beads and the beads are washed. Potential ligand interaction is determined by western blot analysis of the immunoprecipitated complexes with antibody directed towards the epitope tag. An interaction is judged to occur if a band of the anticipated molecular weight of the epitope tagged protein is observed in the

[0475] western blot analysis with a candidate receptor, but is not observed to occur with the other members of the panel of potential receptors.

[0476] Using these assays, the following receptor/ligand interactions have been herein identified:

- (1) PRO10272 binds to PRO5801.



- (2) PRO20110 binds to the human IL-17 receptor (Yao et al., *Cytokine* 9(11):794-800 (1997); also herein designated as PRO1) and to PRO20040.
- (3) PRO10096 binds to PRO20233.
- (4) PRO19670 binds to PRO1890.

[0477] The foregoing written specification is considered to be sufficient to enable one skilled in the art to practice the invention. The present invention is not to be limited in scope by the construct deposited, since the deposited embodiment is intended as a single illustration of certain aspects of the invention and any constructs that are functionally equivalent are within the scope of this invention. The deposit of material herein does not constitute an admission that the written description herein contained is inadequate to enable the practice of any aspect of the invention, including the best mode thereof, nor is it to be construed as limiting the scope of the claims to the specific illustrations that it represents. Indeed, various modifications of the invention in addition to those shown and described herein will become apparent to those skilled in the art from the foregoing description and fall within the scope of the appended claims.

**SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC ACIDS  
ENCODING THE SAME**

Abstract of the Disclosure

The present invention is directed to novel polypeptides and to nucleic acid molecules encoding those polypeptides. Also provided herein are vectors and host cells comprising those nucleic acid sequences, chimeric polypeptide molecules comprising the polypeptides of the present invention fused to heterologous polypeptide sequences, antibodies which bind to the polypeptides of the present invention and to methods for producing the polypeptides of the present invention.

**FIGURE 1**

GGGGCTTCGGCGCCAGCGGCCAGCGCTAGTCGGTCTGGTAAGGATTTACAAAAGGTGCAGGTATG  
AGCAGGTCTGAAGACTAACATTTTGTGAAGTTGTAAACAGAAAACCTGTTAGAAATGTGGTGGT  
TTCAGCAAGGCCTCAGTTTCCTTCCTTCAGCCCTTGTAATTTGGACATCTGCTGCTTTCATATTT  
TCATACATTACTGCAGTAACACTCCACCATATAGACCCGGCTTTACCTTATATCAGTGACACTGG  
TACAGTAGCTCCAGAAAAATGCTTATTTGGGGCAATGCTAAATATTGCGGCAGTTTTATGCATTG  
CTACCATTTATGTTTCGTTATAAGCAAGTTCATGCTCTGAGTCCTGAAGAGAACGTTATCATCAAA  
TTAAACAAGGCTGGCCTTGTAAGTGGAACTGAGTTGTTTAGGACTTTCTATTGTGGCAAACCTT  
CCAGAAAACAACCCTTTTGTGTCACATGTAAGTGGAGCTGTGCTTACCTTTGGTATGGGCTCAT  
TATATATGTTTGTTCAGACCATCCTTTCCTACCAAATGCAGCCCAAAATCCATGGCAAACAAGTC  
TTCTGGATCAGACTGTTGTTGGTTATCTGGTGTGGAGTAAGTGCACTTAGCATGCTGACTTGCTC  
ATCAGTTTTGCACAGTGGCAATTTTGGGACTGATTTAGAACAGAACTCCATTGGAACCCCGAGG  
ACAAAGGTTATGTGCTTCACATGATCACTACTGCAGCAGAATGGTCTATGTCATTTTCCTTCTTT  
GGTTTTTTCCTGACTTACATTCGTGATTTTCAGAAAATTTCTTTACGGGTGGAAGCCAATTTACA  
TGGATTAACCCTCTATGACACTGCACCTTGCCCTATTAACAATGAACGAACACGGCTACTTTCCA  
GAGATATTTGATGAAAGGATAAAATATTTCTGTAATGATTATGATTCTCAGGGATTGGGGAAAGG  
TTCACAGAAGTTGCTTATTCTTCTCTGAAATTTTCAACCACTTAATCAAGGCTGACAGTAACACT  
GATGAATGCTGATAATCAGGAAACATGAAAGAAGCCATTTGATAGATTATTCTAAAGGATATCAT  
CAAGAAGACTATTAAAAACACCTATGCCTATACTTTTTTATCTCAGAAAATAAAGTCAAAGACT  
ATG

## **FIGURE 2**

<subunit 1 of 1, 266 aa, 1 stop

<MW: 29766, pI: 8.39, NX(S/T): 0

MWWFQQGLSFLPSALVIWTSAAFIIFS YITAVTLHHIDPALPYISDTGTVAPEKCLFGAMLNIAAV  
LCIATIYVRYKQVHALSPEENVIIKLNKAGLVLGILSCLGLSIVANFQKTTLFAAHVSGAVLTFG  
MGSLYMFVQTILSYQMOPKIHGKQVFWIRLLLVIWCGVSALSMLTCSSVLHSGNFGTDLEQKLHW  
NPEDKGYVLHMITTAAEWSMSFSFSGFFLT YIRDFQKISLRVEANLHGLTLYDTAPCPINNERTR  
LLSRDI

### **Important features:**

#### **Type II transmembrane domain:**

amino acids 13-33

#### **Other Transmembrane domains:**

amino acids 54-73, 94-113, 160-180, 122-141

#### **N-myristoylation sites.**

amino acids 57-63, 95-101, 99-105, 124-130, 183-189

**FIGURE 3**

CGGACGCGTGGGCGGACGCGTGGGGGAGAGCCGAGTCCCGGCTGCAGCACCTGGGAGAAGGCAG  
 ACCGTGTGAGGGGGCCTGTGGCCCCAGCGTGCTGTGGCCTCGGGGAGTGGGAAGTGGAGGCAGGA  
 GCCTTCCTTACACTTCGCC**ATG**AGTTTCCTCATCGACTCCAGCATCATGATTACCTCCCAGATAC  
 TATTTTTTGGATTTGGGTGGCTTTTCTTCATGCGCCAATTGTTTAAAGACTATGAGATACGTCAG  
 TATGTTGTACAGGTGATCTTCTCCGTGACGTTTGCATTTTCTTGACCATGTTTGAGCTCATCAT  
 CTTTGAAATCTTAGGAGTATTGAATAGCAGCTCCCGTTATTTTCACTGGAAAATGAACCTGTGTG  
 TAATTCTGCTGATCCTGGTTTTTCATGGTGCCTTTTTTACATTGGCTATTTTATTGTGAGCAATATC  
 CGACTACTGCATAAACAACGACTGCTTTTTTCTGTCTCTTATGGCTGACCTTTATGTATTTCTT  
 CTGGAACTAGGAGATCCCTTTCCCATTTCTCAGCCCCAAAACATGGGATCTTATCCATAGAACAGC  
 TCATCAGCCGGGTGGTGTGATTGGAGTGAATCTCATGGCTCTTCTTTCTGGATTTGGTGTCTC  
 AACTGCCCATACACTTACATGTCTTACTTCCTCAGGAATGTGACTGACACGGATATTCTAGCCCT  
 GGAACGGCGACTGCTGCAAACCATGGATATGATCATAAGCAAAAAGAAAAGGATGGCAATGGCAC  
 GGAGAACAATGTTCCAGAAGGGGGGAAGTGCATAACAAACCATCAGGTTTCTGGGGAATGATAAAA  
 AGTGTTACCACTTCAGCATCAGGAAGTGAAAATCTTACTCTTATTCAACAGGAAGTGGATGCTTT  
 GGAAGAATTAAGCAGGCAGCTTTTTCTGGAAACAGCTGATCTATATGCTACCAAGGAGAGAATAG  
 AATACTCCAAAACCTTCAAGGGGAAATATTTTAATTTTCTTGGTTACTTTTTCTCTATTTACTGT  
 GTTTGGAAAATTTTCATGGCTACCATCAATATTGTTTTTGATCGAGTTGGGAAAACGGATCCTGT  
 CACAAGAGGCATTGAGATCACTGTGAATTATCTGGGAATCCAATTTGATGTGAAGTTTTGGTCCC  
 AACACATTTCTTCATTCTTGTGGAATAATCATCGTCACATCCATCAGAGGATTGCTGATCACT  
 CTTACCAAGTTCTTTTTATGCCATCTCTAGCAGTAAGTCCTCCAATGTCATTGTCCTGCTATTAGC  
 ACAGATAATGGGCATGTACTTTGTCTCCTCTGTGCTGCTGATCCGAATGAGTATGCCTTTAGAAT  
 ACCGCACCATAATCACTGAAGTCCTTGGAGAACTGCAGTTCAACTTCTATCACCGTTGGTTTGAT  
 GTGATCTTCCTGGTCAGCGCTCTCTCTAGCATACTCTTCCTCTATTTGGCTCACAAACAGGCACC  
 AGAGAAGCAAATGGCACCT**TGA**ACTTAAGCCTACTACAGACTGTTAGAGGCCAGTGGTTTTCAAAA  
 TTTAGATATAAGAGGGGGGAAAAATGGAACCAGGGCCTGACATTTTATAAACAAACAAAATGCTA  
 TGGTAGCATTTTTTCACCTTCATAGCATACTCCTTCCCCGTCAGGTGATACTATGACCATGAGTAG  
 CATCAGCCAGAACATGAGAGGGGAGAACTAACTCAAGACAATACTCAGCAGAGAGCATCCCGTGTG  
 GATATGAGGCTGGTGTAGAGGCGGAGAGGAGCCAAGAACTAAAGGTGAAAAATACACTGGAAC  
 CTGGGGCAAGACATGTCTATGGTAGCTGAGCCAAACACGTAGGATTTCCGTTTTTAAGGTTACAT  
 GGAAAAGGTTATAGCTTTGCCTTGAGATTGACTCATTAATCAGAGACTGTAACAAAAAAAAAAAA  
 AAAAAAAAAAAGGGCGGCCGCGACTCTAGAGTCGACCTGCAGAAGCTTGGCCGCCATGGCCCAAC  
 TTGTTTATTGCAGCTTATAATG

## **FIGURE 4**

MSFLIDSSIMITSQILFFGFGWLFFMRQLFKDYEIRQYVVQVIFSVTFAFSCTMFELIIFEILGV  
LNSSSRYPFHWMNLCVILLILVFMVPPFYIGYFIVSNIRLLHKQRLLFSCLLWLTFMYFFWKLGDP  
FPILSPKHGILSIEQLISRGVIGVTLMALLSGFGAVNCPYTYMSYFLRNVTDTDILALERRLLQ  
TMDMIISKKKRMAMARRTMFQKGEVHNKPSGFWGMIKSVTTSASGSENLTLIQQEVDAL EELSRQ  
LFLETADLYATKERIEYSKTFKGKYFNFLGYFFSIYCVWKIFMATINIVFDRVGKTDPVTRGIEI  
TVNYLGIQFDVKFWSQHISFILVGIIIVTSIRGLLITLTKFFYAISSSKSSNVIVLLLAQIMGY  
FVSSVLLIRMSMPLEYRTIITEVLGELQFNFYHRWFDVIFLVSALSSILFLYLAHKQAPEKQMAP

**Important features:**

**Signal peptide:**

amino acids 1-23

**Potential transmembrane domains:**

amino acids 37-55, 81-102, 150-168, 288-311, 338-356, 375-398,  
425-444

**N-glycosylation sites.**

amino acids 67-70, 180-183 and 243-246

**Eukaryotic cobalamin-binding proteins**

amino acids 151-160

**FIGURE 5**

AGCAGGGAAATCCGGATGTCTCGGTTATGAAGTGGAGCAGTGAGTGTGAGCCTCAACATAGTTCC  
 AGAACTCTCCATCCGGACTAGTTATTGAGCATCTGCCTCTCATATCACCAGTGGCCATCTGAGGT  
 GTTCCCTGGCTCTGAAGGGGTAGGCACG**ATG**GCCAGGTGCTTCAGCCTGGTGTGCTTCTCACT  
 TCCATCTGGACCACGAGGCTCCTGGTCCAAGGCTCTTTGCGTGCAGAAGAGCTTTCCATCCAGGT  
 GTCATGCAGAATTATGGGGATCACCCCTTGTGAGCAAAAAGGCGAACCAGCAGCTGAATTTACAG  
 AAGCTAAGGAGGCCTGTAGGCTGCTGGGACTAAGTTTGGCCGCAAGGACCAAGTTGAAACAGCC  
 TTGAAAGCTAGCTTTGAAACTTGCAGCTATGGCTGGGTGGAGATGGATTTCGTGGTCATCTCTAG  
 GATTAGCCCAAACCCCAAGTGTGGGAAAAATGGGGTGGGTGTCCTGATTTGGAAGGTTCAGTGA  
 GCCGACAGTTTGCAGCCTATTGTTACAACCTCATCTGATACTTGGACTAACTCGTGCATTCCAGAA  
 ATTATCACCACCAAGATCCCATATTCAACACTCAAACCTGCAACACAAACAACAGAATTTATTGT  
 CAGTGACAGTACCTACTCGGTGGCATCCCCCTTACTCTACAATACCTGCCCCCTACTACTCTCCTC  
 CTGCTCCAGCTTCCACTTCTATTCCACGGAGAAAAAATTGATTTGTGTCACAGAAGTTTTTTATG  
 GAACTAGCACCATGTCTACAGAACTGAACCATTTGTTGAAAATAAAGCAGCATTCAAGAATGA  
 AGCTGCTGGGTTTGGAGGTGTCCCCACGGCTCTGCTAGTGCTTGCTCTCCTCTTCTTTGGTGCTG  
 CAGCTGGTCTTGGATTTTGCTATGTCAAAGGTATGTGAAGGCCTTCCCTTTTACAAACAAGAAT  
 CAGCAGAAGGAAATGATCGAAACCAAAGTAGTAAAGGAGGAGAAGGCCAATGATAGCAACCCATA  
 TGAGGAATCAAAGAAAACCTGATAAAAACCCAGAAGAGTCCAAGAGTCCAAGCAAAACTACCGTGC  
 GATGCCTGGAAGCTGAAGTT**TAG**ATGAGACAGAAATGAGGAGACACACCTGAGGCTGGTTTTCTTT  
 CATGCTCCTTACCCTGCCCCAGCTGGGGAAATCAAAGGGCCAAAGAACCAAGAAGAAAGTCCA  
 CCCTTGGTTCCCTAACTGGAATCAGCTCAGGACTGCCATTGGACTATGGAGTGCACCAAAGAGAAT  
 GCCCTTCTCCTTATTGTAACCCCTGTCTGGATCCTATCCTCCTACCTCCAAAGCTTCCCACGGCCT  
 TTCTAGCCTGGCTATGTCCTAATAATATCCCACCTGGGAGAAAGGAGTTTTGCAAAGTGCAAGGAC  
 CTAACATCTCATCAGTATCCAGTGGTAAAAAGGCCTCCTGGCTGTCTGAGGCTAGGTGGGTG  
 AAAGCCAAGGAGTCACTGAGACCAAGGCTTTCTCTACTGATTCCGCAGCTCAGACCCTTTCTTCA  
 GCTCTGAAAGAGAAACACGTATCCACCTGACATGTCTTCTGAGCCCGGTAAGAGCAAAAGAAT  
 GGCAGAAAAGTTTAGCCCCCTGAAAGCCATGGAGATTCTCATAACTTGAGACCTAATCTCTGTAAA  
 GCTAAAATAAAGAAATAGAACAAGGCTGAGGATACGACAGTACACTGTCAGCAGGGACTGTAAAC  
 ACAGACAGGGTCAAAGTGTTTTCTCTGAACACATTGAGTTGGAATCACTGTTTAGAACACACACA  
 CTTACTTTTTCTGGTCTCTACCACTGCTGATATTTCTCTAGGAAATATACTTTTACAAGTAACA  
 AAAATAAAAACCTCTTATAAATTTCTATTTTTATCTGAGTTACAGAAATGATTACTAAGGAAGATT  
 ACTCAGTAATTTGTTTAAAAAGTAATAAAATTCAACAAACATTTGCTGAATAGCTACTATATGTC  
 AAGTGCTGTGCAAGGTATTACACTCTGTAATTGAATATTATTCCTCAAAAAATTGCACATAGTAG  
 AACGCTATCTGGGAAGCTATTTTTTTTCAGTTTTGATATTTCTAGCTTATCTACTTCCAAACTAAT  
 TTTTATTTTTGCTGAGACTAATCTTATTCATTTTCTCTAATATGGCAACCATTATAACCTTAATT  
 TATTATTAACATACTAAGAAGTACATTGTTACCTCTATATACCAAAGCACATTTTAAAAGTGCC  
 ATTAACAAATGTATCACTAGCCCTCCTTTTTTCCAACAAGAAGGGACTGAGAGATGCAGAAATATT  
 TGTGACAAAAAATTAAAGCATTTAGAAAACCTT

## **FIGURE 6**

MARCFSLVLLLLTSIWTTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLLG  
LSLAGKDQVETALKASFETCSYGWVGDFVVISRISPNPKCGKNGVGVLWKVPVSRQFAAYCYN  
SSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPAPTTTPAPASTSIPR  
RKKLICVTEVFMETSTMSTETEPFVENKAAFKNAAAGFGGVPTALLVLALLFFGAAAGLGFCYVK  
RYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKTDKNPEESKSPSKTTVRCLEAEV

**Signal sequence:**

amino acids 1-16

**Transmembrane domain:**

amino acids 235-254

**N-glycosylation site.**

amino acids 53-57, 130-134, 289-293

**Casein kinase II phosphorylation site.**

amino acids 145-149, 214-218

**Tyrosine kinase phosphorylation site.**

amino acids 79-88

**N-myristoylation site.**

amino acids 23-29, 65-71, 234-240, 235-239, 249-255, 253-259



**FIGURE 7**

CGCCGCGCTCCCGCACCCGCGGCCCGCCACCGCGCCGCTCCCGCATCTGCACCCGCAGCCCGGC  
GGCCTCCCGGCGGGAGCGAGCAGATCCAGTCCGGCCCCGCAGCGCAACTCGGTCCAGTCGGGGCGG  
CGGCTGCGGGCGCAGAGCGGAG**ATG**CAGCGGCTTGGGGCCACCCTGCTGTGCCTGCTGCTGGCGG  
CGGCGGTCCCCACGGCCCCCGCGCCCGCTCCGACGGCGACCTCGGCTCCAGTCAAGCCCGGCCCG  
GCTCTCAGCTACCCGCAGGAGGAGGCCACCCTCAATGAGATGTTCCGCGAGGTTGAGGAACTGAT  
GGAGGACACGCAGCACAAATTGCGCAGCGCGGTGGAAGAGATGGAGGCAGAAGAAGCTGCTGCTA  
AAGCATCATCAGAAGTGAACCTGGCAAACCTTACCTCCCAGCTATCACAATGAGACCAACACAGAC  
ACGAAGGTTGGAAATAATACCATCCATGTGCACCGAGAAATTCACAAGATAACCAACAACCAGAC  
TGGACAAATGGTCTTTTCAGAGACAGTTATCACATCTGTGGGAGACGAAGAAGGCAGAAGGAGCC  
ACGAGTGCATCATCGACGAGGACTGTGGGCCAGCATGTACTGCCAGTTTGCCAGCTTCCAGTAC  
ACCTGCCAGCCATGCCGGGGCCAGAGGATGCTCTGCACCCGGGACAGTGAGTGCTGTGGAGACCA  
GCTGTGTGTCTGGGGTCACTGCACCAAAATGGCCACCAGGGGCAGCAATGGGACCATCTGTGACA  
ACCAGAGGGACTGCCAGCCGGGGCTGTGCTGTGCCTTCCAGAGAGGCCTGCTGTTCCCTGTGTGC  
ACACCCCTGCCCGTGGAGGGCGAGCTTTGCCATGACCCGCCAGCCGGCTTCTGGACCTCATCAC  
CTGGGAGCTAGAGCCTGATGGAGCCTTGGACCGATGCCCTTGTGCCAGTGGCCTCCTCTGCCAGC  
CCCACAGCCACAGCCTGGTGTATGTGTGCAAGCCGACCTTCGTGGGGAGCCGTGACCAAGATGGG  
GAGATCCTGCTGCCCAGAGAGGTCCCCGATGAGTATGAAGTTGGCAGCTTCATGGAGGAGGTGCG  
CCAGGAGCTGGAGGACCTGGAGAGGAGCCTGACTGAAGAGATGGCGCTGGGGGAGCCTGCGGCTG  
CCGCCGCTGCACTGCTGGGAGGGGAAGAGATT**TAGAT**CTGGACCAGGCTGTGGGTAGATGTGCAA  
TAGAAATAGCTAATTTATTTCCCCAGGTGTGTGCTTTAGGCGTGGGCTGACCAGGCTTCTTCCTA  
CATCTTCTTCCAGTAAGTTTCCCCTCTGGCTTGACAGCATGAGGTGTTGTGCATTTGTTTCAGCT  
CCCCCAGGCTGTTCTCCAGGCTTCACAGTCTGGTGCTTGGGAGAGTCAGGCAGGGTTAAACTGCA  
GGAGCAGTTTGGCACCCCTGTCCAGATTATTGGCTGCTTTGCCTCTACCAGTTGGCAGACAGCCG  
TTTGTCTACATGGCTTTGATAATTGTTTGGAGGGGAGGAGATGGAAACAATGTGGAGTCTCCCTC  
TGATTGGTTTTTGGGGAAATGTGGAGAAGAGTGCCCTGCTTTGCAAACATCAACCTGGCAAAAATG  
CAACAAATGAATTTTCCACGCAGTTCTTTCCATGGGCATAGGTAAGCTGTGCCTTCAGCTGTTGC  
AGATGAAATGTTCTGTTTACCCTGCATTACATGTGTTTATTTCATCCAGCAGTGTTGCTCAGCTCC  
TACCTCTGTGCCAGGGCAGCATTTTCATATCCAAGATCAATTCCCTCTCTCAGCACAGCCTGGGG  
AGGGGGTCATTGTTCTCCTCGTCCATCAGGGATCTCAGAGGCTCAGAGACTGCAAGCTGCTTGCC  
CAAGTCACACAGCTAGTGAAGACCAGAGCAGTTTCATCTGGTTGTGACTCTAAGCTCAGTGCTCT  
CTCCACTACCCACACCAGCCTTGGTGCCACCAAAAGTGCTCCCCAAAAGGAAGGAGAATGGGAT  
TTTTCTTGAGGCATGCACATCTGGAATTAAGGTCAAACCTAATTCTCACATCCCTCTAAAAGTAAA  
CTACTGTTAGGAACAGCAGTGTTCTCACAGTGTGGGGCAGCCGTCCTTCTAATGAAGACAATGAT  
ATTGACACTGTCCCTCTTTGGCAGTTGCATTAGTAACCTTTGAAAGGTATATGACTGAGCGTAGCA  
TACAGGTTAACCTGCAGAAACAGTACTTAGGTAATTGTAGGGCGAGGATTATAAATGAAATTTGC  
AAAATCACTTAGCAGCAACTGAAGACAATTATCAACCACGTGGAGAAAATCAAACCGAGCAGGGC  
TGTGTGAAACATGGTTGTAATATGCGACTGCGAACACTGAACTCTACGCCACTCCACAAATGATG  
TTTTCAGGTGTCATGGACTGTTGCCACCATGTATTTCATCCAGAGTTCTTAAAGTTTAAAGTTGCA  
CATGATTGTATAAGCATGCTTTCTTTGAGTTTTTAAATTATGTATAAACATAAGTTGCATTTAGAA  
ATCAAGCATAAATCACTTCAACTGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

## **FIGURE 8**

MQRLGATLLCLLLAAVPTAPAPAPTATSAPVKPGPALSYPQEEATLNEMFREVEELMEDTQHKL  
RSAVEEMEAEAAAKASSEVNLANLPPSYHNETNTDTKVGNNTIHVHREIHKITNNQTGQMVFESE  
TVITSVGDEEGRRSHECIIDEDCGPSMYCQFASFQYTCQPCRGQRMMLCTRDSECCGDQLCVWGHC  
TKMATRGSNGTICDNQRDCQPGLCCAFQRGLLFPVCTPLPVEGELCHDPASRLDLITWELEPDG  
ALDRCPCASGLLCQPHSHSLVYVCKPTFVGSRDQDGEILLPREVPDEYEVGSGFMEEVRQELEDLE  
RSLTEEMALGEPAAAAAALLGGEEI

**Signal sequence:**

amino acids 1-19

**N-glycosylation site.**

amino acids 96-100, 106-110, 121-125, 204-208

**Casein kinase II phosphorylation site.**

amino acids 46-50, 67-71, 98-102, 135-139, 206-210, 312-316, 327-331

**N-myristoylation site.**

amino acids 202-208, 217-223

**Amidation site.**

amino acids 140-144

**FIGURE 9**

**CGG**ACGCGTGGGCGGACGCGTGGGGGCTGTGAGAAAGTGCCAATAAATACATCATGCAACCCAC  
GGCCACCTTGTGAACTCCTCGTGCCAGGGCTGATGTGCGTCTTCCAGGGCTACTCATCCAAAG  
GCCTAATCCAACGTTCTGTCTTCAATCTGCAAATCTATGGGGTCCTGGGGCTCTTCTGGACCCCT  
AACTGGGTACTGGCCCTGGGCCAATGCGTCCTCGCTGGAGCCTTTGCCTCCTTCTACTGGGCCTT  
CCACAAGCCCCAGGACATCCCTACCTTCCCCTTAATCTCTGCCTTCATCCGCACACTCCGTTACC  
ACACTGGGTCATTGGCATTGAGGCCCTCATCCTGACCCTTGTGCAGATAGCCCGGGTCATCTTG  
GAGTATATTGACCACAAGCTCAGAGGAGTGCAGAACCCTGTAGCCCGCTGCATCATGTGCTGTTT  
CAAGTGCTGCCTCTGGTGTCTGGAAAATTTATCAAGTTCCTAAACCGCAATGCATACATCATGA  
TCGCCATCTACGGGAAGAATTTCTGTGTCTCAGCCAAAAATGCGTTCATGCTACTCATGCGAAAC  
ATTGTCAGGGTGGTCGTCCTGGACAAAGTCACAGACCTGCTGCTGTTCTTTGGGAAGCTGCTGGT  
GGTCGGAGGCGTGGGGGTCCTGTCCTTCTTTTTTTCTCCGGTCGCATCCCGGGGCTGGGTAAAG  
ACTTTAAGAGCCCCACCTCAACTATTACTGGCTGCCCATCATGACCTCCATCCTGGGGGCCTAT  
GTCATCGCCAGCGGCTTCTTCAGCGTTTTTCGGCATGTGTGTGGACACGCTCTTCCTCTGCTTCCT  
GGAAGACCTGGAGCGGAACAACGGCTCCCTGGACCGGCCCTACTACATGTCCAAGAGCCTTCTAA  
AGATTCTGGGCAAGAAGAACGAGGCGCCCCCGGACAACAAGAAGAGGAAGAAG**TGA**CAGCTCCGG  
CCCTGATCCAGGACTGCACCCACCCACCGTCCAGCCATCCAACCTCACTTCGCCTTACAGGT  
CTCCATTTTGTGGTAAAAAAGGTTTTAGGCCAGGCGCCGTGGCTCACGCCTGTAATCCAACACT  
TTGAGAGGCTGAGGCGGGCGGATCACCTGAGTCAGGAGTTCGAGACCAGCCTGGCCAACATGGTG  
AAACCTCCGTCTCTATTAAAAATACAAAAATTAGCCGAGAGTGGTGGCATGCACCTGTCATCCCA  
GCTACTCGGGAGGCTGAGGCAGGAGAATCGCTTGAACCCGGGAGGCAGAGGTTGCAGTGAGCCGA  
GATCGCGCCACTGCACTCCAACCTGGGTGACAGACTCTGTCTCCAAAACAAAACAAACAAA  
AAGATTTTATTAAAGATATTTTGTTAACTC

## **FIGURE 10**

RTRGRTRGGCEKVPINTSCNPTAHLVNSSCPGLMCVFQGYSSKGLIQRSVFNLQIYGVLGLFWTL  
NWVLALGQCVLGAFASFYWAFHKPQDIPTFPLISAFIRTLRYHTGSLAFGALILTLVQIARVIL  
EYIDHKLRGVQNPVARCIMCCFKCCLWCLEKFIKFLNRNAYIMIAIYGKNFCVSAKNAFMLLMRN  
IVRVVVLDDKVTDLLLFFGKLLVVGGVGVLSFFFFSGRIPGLGKDFKSPHLNYYWLPIMTSILGAY  
VIASGFFSVFGMCVDTLFLCFLEDLERNNGSLDRPYMSKSLKILGKKNEAPPDNKKRKK

**Important features:**

**Transmembrane domains:**

amino acids 57-80 (type II), 110-126, 215-231, 254-274

**N-glycosylation sites.**

amino acids 16-20, 27-31, 289-293

**Hypothetical YBR002c family proteins.**

amino acids 276-288

**Ammonium transporters proteins.**

amino acids 204-231

**N-myristoylation sites.**

amino acids 60-66, 78-84

**Amidation site.**

amino acids 306-310

**FIGURE 11**

GCCCCGCGCCCGGCGCCGGGCGCCCGAAGCCGGGAGCCACCGCCATGGGGGCCTGCCTGGGAGCC  
TGCTCCCTGCTCAGCTGCGCGTCCTGCCTCTGCGGCTCTGCCCCCTGCATCCTGTGCAGCTGCTG  
CCCCGCCAGCCGCAACTCCACCGTGAGCCGCCTCATCTTCACGTTCTTCCTCTTCCTGGGGGTGC  
TGGTGTCCATCATTATGCTGAGCCCGGGCGTGAGAGTCAGCTCTACAAGCTGCCCTGGGTGTGT  
GAGGAGGGGGCCGGGATCCCCACCGTCCTGCAGGGCCACATCGACTGTGGCTCCCTGCTTGGCTA  
CCGCGCTGTCTACCGCATGTGCTTCGCCACGGCGGCCTTCTTCTTCTTTTTCACCCTGCTCA  
TGCTCTGCGTGAGCAGCAGCCGGGACCCCCGGGCTGCCATCCAGAATGGGTTTTGGTTCTTTAAG  
TTCCTGATCCTGGTGGGCCTCACCGTGGGTGCCTTCTACATCCCTGACGGCTCCTTCACCAACAT  
CTGGTTCTACTTCGGCGTCGTGGGCTCCTTCCTCTTCATCCTCATCCAGCTGGTGCTGCTCATCG  
ACTTTGCGCACTCCTGGAACCAGCGGTGGCTGGGCAAGGCCGAGGAGTGCATTCCTGTGCCTGG  
TACGCAGGCCTCTTCTTCTTCACTCTCCTCTTCTACTTGCTGTGATCGCGGCCGTGGCGCTGAT  
GTTTCATGTACTACACTGAGCCCAGCGGCTGCCACGAGGGCAAGGTCTTCATCAGCCTCAACCTCA  
CCTTCTGTGTCTGCGTGTCCATCGCTGCTGTCTGCCCCAAGGTCCAGGACGCCAGCCCAACTCG  
GGTCTGCTGCAGGCCTCGGTCATCACCTCTACACCATGTTTGTACCTGGTCAGCCCTATCCAG  
TATCCCTGAACAGAAATGCAACCCCCATTTGCCAACCAGCTGGGCAACGAGACAGTTGTGGCAG  
GCCCCGAGGGCTATGAGACCCAGTGGTGGGATGCCCCGAGCATTGTGGGCCTCATCATCTTCCTC  
CTGTGCACCCTCTTCATCAGTCTGCGCTCCTCAGACCACCGGCAGGTGAACAGCCTGATGCAGAC  
CGAGGAGTGCCACCTATGCTAGACGCCACACAGCAGCAGCAGCAGCAGGTGGCAGCCTGTGAGG  
GCCGGGCCTTTGACAACGAGCAGGACGGCGTCACCTACAGCTACTCCTTCTTCCACTTCTGCCTG  
GTGCTGGCCTCACTGCACGTCATGATGACGCTCACCAACTGGTACAAGCCCGGTGAGACCCGGAA  
GATGATCAGCACGTGGACCGCCGTGTGGGTGAAGATCTGTGCCAGCTGGGCAGGGCTGCTCCTCT  
ACCTGTGGACCCTGGTAGCCCCACTCCTCCTGCGCAACCGCGACTTCAGCTTGAGGCAGCCTCACA  
GCCTGCCATCTGGTGCCTCCTGCCACCTGGTGCCTCTCGGCTCGGTGACAGCCAACCTGCCCCCT  
CCCCACACCAATCAGCCAGGCTGAGCCCCCACCCTGCCCCAGCTCCAGGACCTGCCCCTGAGCC  
GGGCCTTCTAGTCGTAGTGCCTTCAGGGTCCGAGGAGCATCAGGCTCCTGCAGAGCCCCATCCCC  
CCGCCACACCCACACGGTGGAGCTGCCTCTTCCTTCCCCCTCCTCCCTGTTGCCATACTCAGCAT  
CTCGGATGAAAGGGCTCCCTTGTCTCAGGCTCCACGGGAGCGGGGCTGCTGGAGAGAGCGGGGA  
ACTCCCACACAGTGGGGCATCCGGCACTGAAGCCCTGGTGTTCCTGGTCACGTCCCCCAGGGGA  
CCCTGCCCCCTTCTGGACTTCGTGCCTTACTGAGTCTCTAAGACTTTTTCTAATAAACAAGCCA  
GTGCGTGTAACAAAAA

## **FIGURE 12**

MGACLGACSLSCASCLCGSAPCILCSCCPASRNSTVSRLIFTFFLFLGVLVSIIMLSPGVESQL  
YKLPWVCEEGAGIPTVLQGHIDCGSLLGYRAVYRMCFATAAFFFFFFFFTLLMLCVSSSRDPRAAIQ  
NGFWFFKFLILVGLTVGAFYIPDGSFTNIWFYFGVVGSFLFILIQLVLLIDFAHSWNQRWLKAE  
ECDSRAWYAGLFFFTLLFYLLSIAAVALMFMYYTEPSGCHEGKVFISLNLTFVCVVSIAAVLPKV  
QDAQPNSGLLQASVITLYTMFVTWSALSSIPEQKCNPHLPTQLGNETVVAGPEGYETQWWDA PSI  
VGLIIFLLCTLFISLRSSDHRQVNSLMQTEECPPMLDATQQQQQVAACEGRAFDNEQDGV TYSY  
SFFHFCLVLASLHVMMTLTNWYKPGETRKMISTWTAVWVKICASWAGLLLYLWTLVAPLLLRNRD  
FS

**Signal sequence:**

amino acids 1-20

**Transmembrane domains:**

amino acids 40-58, 101-116, 134-150, 162-178, 206-223, 240-257,  
272-283, 324-340, 391-406, 428-444

**FIGURE 13**

CGGGCCAGCCTGGGGCGGGCCGGCCAGGAACCACCCGTTAAGGTGTCTTCTCTTTAGGGATGGTGA  
GGTTGGAAAAAGACTCCTGTAACCCTCCTCCAGGATGAACCACCTGCCAGAAGACATGGAGAACG  
CTCTCACCGGGAGCCAGAGCTCCCATGCTTCTCTGCGCAATATCCATTCCATCAACCCACACAA  
CTCATGGCCAGGATTGAGTCCTATGAAGGAAGGGAAAAGAAAGGCATATCTGATGTCAGGAGGAC  
TTTCTGTTTGTTTGTCACCTTTGACCTCTTATTCGTAACATTACTGTGGATAATAGAGTTAAATG  
TGAATGGAGGCATTGAGAACACATTAGAGAAGGAGGTGATGCAGTATGACTACTATTCTTCATAT  
TTTGATATATTTCTTCTGGCAGTTTTTTCGATTTAAAGTGTTAATACTTGTCATATGCTGTGTGCAG  
ACTGCGCCATTGGTGGGCAATAGCGTTGACAACGGCAGTGACCAGTGCCTTTTTACTAGCAAAAG  
TGATCCTTTCGAAGCTTTTCTCTCAAGGGGCTTTTGGCTATGTGCTGCCCATCATTTCATTCATC  
CTTGCCTGGATTGAGACGTGGTTCCTGGATTTCAAAGTGTTACCTCAAGAAGCAGAAGAAGAAAA  
CAGACTCCTGATAGTTCAGGATGCTTCAGAGAGGGCAGCACTTATACCTGGTGGTCTTTCTGATG  
GTCAGTTTTATTCCCCTCCTGAATCCGAAGCAGGATCTGAAGAAGCTGAAGAAAAACAGGACAGT  
GAGAAACCACTTTTAGAACTATGAGTACTACTTTTGTTAAATGTGAAAAACCTCACAGAAAGTC  
ATCGAGGCAAAAAGAGGCAGGCAGTGGAGTCTCCCTGTGACAGTAAAGTTGAAATGGTGACGTC  
CACTGCTGGCTTTATTGAACAGCTAATAAAGATTTATTTATTGTAATACCTCACAAACGTTGTAC  
CATATCCATGCACATTTAGTTGCCTGCCTGTGGCTGGTAAGGTAATGTCATGATTCATCCTCTCT  
TCAGTGAGACTGAGCCTGATGTGTTAACAAATAGGTGAAGAAAGTCTTGCTGTATTCTTAATC  
AAAAGACTTAATATATTGAAGTAACACTTTTTTAGTAAGCAAGATACCTTTTTTATTTCAATTCAC  
AGAATGGAATTTTTTTGTTTCATGTCTCAGATTTATTTTGTATTTCTTTTTTAACACTCTACATT  
TCCCTTGTTTTTTAACTCATGCACATGTGCTCTTTGTACAGTTTTAAAAAGTGTAATAAAATCTG  
ACATGTCAATGTGGCTAGTTTTATTTTTCTTGTTTTGCATTATGTGTATGGCCTGAAGTGTTGGA  
CTTGCAAAAGGGGAAGAAAGGAATTGCGAATACATGTAAAATGTCACCAGACATTTGTATTATTT  
TTATCATGAAATCATGTTTTCTCTGATTGTTCTGAAATGTTCTAAATACTCTTATTTTGAATGC  
ACAAAATGACTTAAACCATTCATATCATGTTTCCTTTGCGTTCAGCCAATTTCAATTAAAATGAA  
CTAAATTAAAAA

## **FIGURE 14**

MNHLPEDMENALTGSQSSHASLRNIHSINPTQLMARIESYEGREKKGISDVRRTFCLFVTFDLLF  
VTLLWIIELNVNGGIENLEKEVMQYDYYSSYFDIFLLAVFRFKVLILAYAVCRLRHWWAIALTT  
AVTSAFLLAKVILSKLFSQGAFGYVLPPIISFILAWIETWFLDFKVLPPQEAEENRLLIVQDASER  
AALIPGGLSDGQFYSPPESEAGSEEAEEKQDSEKPLLEL

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-20

#### **Transmembrane domains:**

amino acids 54-72, 100-118, 130-144, 146-166

#### **N-myristoylation sites.**

amino acids 14-20, 78-84, 79-85, 202-208, 217-223



**FIGURE 15**

ACTCGAACGCAGTTGCTTCGGGACCCAGGACCCCCTCGGGCCCCGACCCGCCAGGAAAGACTGAGG  
 CCGCGGCCTGCCCCGCCCCGGCTCCCTGCGCCGCCGCCGCTCCCGGGACAGAAG**ATG**TGCTCCAG  
 GGTCCCTCTGCTGCTGCCGCTGCTCCTGCTACTGGCCCTGGGGCCTGGGGTGACAGGGCTGCCCCAT  
 CCGGCTGCCAGTGACAGCCAGCCACAGACAGTCTTCTGCACTGCCCCGCCAGGGGACCACGGTGCCC  
 CGAGACGTGCCACCCGACACGGTGGGGCTGTACGTCTTTGAGAACGGCATCACCATGCTCGACGC  
 AGGCAGCTTTGCCGGCCTGCCGGGCCTGCAGCTCCTGGACCTGTCACAGAACCAGATCGCCAGCC  
 TGCCAGCGGGGTCTTCCAGCCACTCGCCAACCTCAGCAACCTGGACCTGACGGCCAACAGGCTG  
 CATGAAATCACCAATGAGACCTTCCGTGGCCTGCGGCGCCTCGAGCGCCTCTACCTGGGCAAGAA  
 CCGCATCCGCCACATCCAGCCTGGTGCCTTCGACACGCTCGACCGCCTCCTGGAGCTCAAGCTGC  
 AGGACAACGAGCTGCGGGCACTGCCCCCGCTGCGCCTGCCCCGCCTGCTGCTGCTGGACCTCAGC  
 CACAACAGCCTCCTGGCCCTGGAGCCCGGCATCCTGGACACTGCCAACGTGGAGGCGCTGCGGCT  
 GGCTGGTCTGGGGCTGCAGCAGCTGGACGAGGGGCTCTTCAGCCGCTTGCGCAACCTCCACGACC  
 TGGATGTGTCCGACAACCAGCTGGAGCGAGTGCCACCTGTGATCCGAGGCCTCCGGGGCCTGACG  
 CGCCTGCGGCTGGCCGGCAACACCCGCATTGCCAGCTGCGGCCCGAGGACCTGGCCGGCCTGGC  
 TGCCCTGCAGGAGCTGGATGTGAGCAACCTAAGCCTGCAGGCCCTGCCTGGCGACCTCTCGGGCC  
 TCTTCCCCCGCCTGCGGCTGCTGGCAGCTGCCCCGAACCCCTTCAACTGCGTGTGCCCCCTGAGC  
 TGGTTTGGCCCCCTGGGTGCGCGAGAGCCACGTACACTGGCCAGCCCTGAGGAGACGCGCTGCCA  
 CTTCCCCGCCAAGAACGCTGGCCGGCTGCTCCTGGAGCTTGACTACGCCGACTTTGGCTGCCCAG  
 CCACCACCACACAGCCACAGTGCCCACCACGAGGCCCGTGGTGCGGGAGCCCACAGCCTTGTCT  
 TCTAGCTTGGCTCCTACCTGGCTTAGCCCCACAGCGCCGGCCACTGAGGCCCCCAGCCCGCCCTC  
 CACTGCCCCACCGACTGTAGGGCCTGTCCCCAGCCCCAGGACTGCCACCGTCCACCTGCCTCA  
 ATGGGGGCACATGCCACCTGGGGACACGGCACACCTGGCGTGCTTGTGCCCCGAAGGCTTCACG  
 GGCCTGTACTGTGAGAGCCAGATGGGGCAGGGGACACGGCCCCAGCCCTACACCAGTCACGCCGAG  
 GCCACCACGGTCCCTGACCCTGGGCATCGAGCCGGTGAGCCCCACCTCCCTGCGCGTGGGGCTGC  
 AGCGCTACCTCCAGGGGAGCTCCGTGCAGCTCAGGAGCCTCCGTCTCACCTATCGCAACCTATCG  
 GGCCCTGATAAGCGGCTGGTGACGCTGCGACTGCCTGCCTCGCTCGCTGAGTACAGGTCAACCA  
 GCTGCGGCCCAACGCCACTTACTCCGTCTGTGTATGCCTTTGGGGCCCGGGCGGGTGCCGGAGG  
 GCGAGGAGGCCTGCGGGGAGGCCATAACCCCCAGCCGTCCACTCCAACCACGCCCCAGTCACC  
 CAGGCCCGCGAGGGCAACCTGCCGCTCCTCATTGCGCCCGCCCTGGCCGCGGTGCTCCTGGCCGC  
 GCTGGCTGCGGTGGGGGCAGCCTACTGTGTGCGGCGGGGGCGGGCCATGGCAGCAGCGGCTCAGG  
 ACAAGGGCAGGTGGGGCCAGGGGCTGGGCCCTGGAACCTGGAGGGAGTGAAGGTCCCCTTGAG  
 CCAGGCCCGAAGGCAACAGAGGGCGGTGGAGAGGCCCTGCCCAGCGGGTCTGAGTGTGAGGTGCC  
 ACTCATGGGCTTCCCAGGGCCTGGCCTCCAGTCACCCCTCCACGCAAAGCCCTACATC**TAAG**CCA  
 GAGAGAGACAGGGCAGCTGGGGCCGGGCTCTCAGCCAGTGAGATGGCCAGCCCCCTCCTGCTGCC  
 ACACCACGTAAAGTTCTCAGTCCCAACCTCGGGGATGTGTGCAGACAGGGCTGTGTGACCACAGCT  
 GGGCCCTGTTCCCTCTGGACCTCGGTCTCCTCATCTGTGAGATGCTGTGGCCCAGCTGACGAGCC  
 CTAACGTCCCCAGAACCGAGTGCCATATGAGGACAGTGTCCGCCCTGCCCTCCGCAACGTGCAGTC  
 CCTGGGCACGGCGGGGCCCTGCCATGTGCTGGTAACGCATGCCTGGGTCTGCTGGGCTCTCCAC  
 TCCAGGCGGACCCCTGGGGGCCAGTGAAGGAAGCTCCCGGAAAGAGCAGAGGGAGAGCGGGTAGGC  
 GGCTGTGTGACTCTAGTCTTGGCCCCAGGAAGCGAAGGAACAAAGAAACTGGAAAGGAAGATGC  
 TTTAGGAACATGTTTTTGTCTTTTTTAAATATATATATTTATAAGAGATCCTTTCCCATTTATTCT  
 GGAAGATGTTTTTCAAACCTCAGAGACAAGGACTTTGGTTTTTTGTAAGACAAACGATGATATGAA  
 GGCCTTTTGTAAAGAAAAAATAAAAGATGAAGTGTGAAA

## **FIGURE 16**

MCSRVP LLLPL LLL L L L L L A L G P G V Q G C P S G C Q C S Q P Q T V F C T A R Q G T T V P R D V P P D T V G L Y V F E N G I T  
 M L D A G S F A G L P G L Q L L D L S Q N Q I A S L P S G V F Q P L A N L S N L D L T A N R L H E I T N E T F R G L R R L E R L Y  
 L G K N R I R H I Q P G A F D T L D R L L E L K L Q D N E L R A L P P L R L P R L L L D L S H N S L L A L E P G I L D T A N V E  
 A L R L A G L G L Q Q L D E G L F S R L R N L H D L D V S D N Q L E R V P P V I R G L R G L T R L R L A G N T R I A Q L R P E D L  
 A G L A A L Q E L D V S N L S L Q A L P G D L S G L F P R L R L L A A A R N P F N C V C P L S W F G P W V R E S H V T L A S P E E  
 T R C H F P P K N A G R L L L E L D Y A D F G C P A T T T T A T V P T T R P V V R E P T A L S S S L A P T W L S P T A P A T E A P  
 S P P S T A P P T V G P V P Q P Q D C P P S T C L N G G T C H L G T R H H L A C L C P E G F T G L Y C E S Q M G Q G T R P S P T P  
 V T P R P P R S L T L G I E P V S P T S L R V G L Q R Y L Q G S S V Q L R S L R L T Y R N L S G P D K R L V T L R L P A S L A E Y  
 T V T Q L R P N A T Y S V C V M P L G P G R V P E G E E A C G E A H T P P A V H S N H A P V T Q A R E G N L P L L I A P A L A A V  
 L L A A L A A V G A A Y C V R R G R A M A A A A Q D K G Q V G P G A G P L E L E G V K V P L E P G P K A T E G G G E A L P S G S E  
 C E V P L M G F P G P G L Q S P L H A K P Y I

### **Important features:**

#### **Signal peptide:**

amino acids 1-23

#### **Transmembrane domain:**

amino acids 579-599

#### **EGF-like domain cysteine pattern signature.**

amino acids 430-442

#### **Leucine zipper pattern.**

amino acids 197-219, 269-291

#### **N-glycosylation sites.**

amino acids 101-105, 117-121, 273-277, 500-504, 528-532

#### **Tyrosine kinase phosphorylation sites.**

amino acids 124-131, 337-345

#### **N-myristoylation sites.**

amino acids 23-29, 27-33, 70-76, 142-148, 187-193, 348-354, 594-600,  
 640-646

**FIGURE 17**

GCAGCGGCGAGGCGGCGGTGGTGGCTGAGTCCGTGGTGGCAGAGGCGAAGGCGACAGCTC**ATGCG**  
GGTCCGGATAGGGCTGACGCTGCTGCTGTGTGCGGTGCTGCTGAGCTTGGCCTCGGCGTCCCTCGG  
ATGAAGAAGGCAGCCAGGATGAATCCTTAGATTCCAAGACTACTTTGACATCAGATGAGTCAGTA  
AAGGACCATACTACTGCAGGCAGAGTAGTTGCTGGTCAAATATTTCTTGATTGAGAAGAATCTGA  
ATTAGAATCCTCTATTCAAGAAGAGGAAGACAGCCTCAAGAGCCAAGAGGGGGAAAGTGTCACAG  
AAGATATCAGCTTTCTAGAGTCTCCAAATCCAGAAAACAAGGACTATGAAGAGCCAAAGAAAGTA  
CGGAAACCAGCTTTGACCGCCATTGAAGGCACAGCACATGGGGAGCCCTGCCACTTCCCTTTTCT  
TTTCCTAGATAAGGAGTATGATGAATGTACATCAGATGGGAGGGAAGATGGCAGACTGTGGTGTG  
CTACAACCTATGACTACAAAGCAGATGAAAAGTGGGGCTTTTGTGAACTGAAGAAGAGGCTGCT  
AAGAGACGGCAGATGCAGGAAGCAGAAATGATGTATCAAACCTGGAATGAAAATCCTTAATGGAAG  
CAATAAGAAAAGCCAAAAAAGAGAAGCATATCGGTATCTCCAAAAGGCAGCAAGCATGAACCATA  
CCAAAGCCCTGGAGAGAGTGTGCATATGCTCTTTTATTTGGTGATTACTTGCCACAGAATATCCAG  
GCAGCGAGAGAGATGTTTGAGAAGCTGACTGAGGAAGGCTCTCCAAGGGACAGACTGCTCTTGG  
CTTTCTGTATGCCTCTGGACTTGGTGTTAATTCAAGTCAGGCAAAGGCTCTTGTATATTATACAT  
TTGGAGCTCTTGGGGGCAATCTAATAGCCACATGGTTTTGGTAAGTAGACTT**TAGT**GGAAGGCT  
AATAATATTAACATCAGAAGAATTTGTGGTTTATAGCGGCCACAACCTTTTTCAGCTTTCATGATC  
CAGATTTGCTTGTATTAAGACCAAATATTCAGTTGAACTTCCTTCAAATTCCTGTTAATGGATAT  
AACACATGGAATCTACATGTAAATGAAAGTTGGTGGAGTCCACAATTTTCTTTAAATGATTAG  
TTTGGCTGATTGCCCCATAAAAGAGAGATCTGATAAATGGCTCTTTTAAATTTTCTCTGAGTTG  
GAATTGTCAGAATCATTTTTTTACATTAGATTATCATAATTTTAAAAATTTTCTTTAGTTTTTCA  
AAATTTTGTAATGGTGGCTATAGAAAAACAACATGAAATATTATACAATATTTTGCAACAATGC  
CCTAAGAATTGTAAAATTCATGGAGTTATTTGTGCAGAATGACTCCAGAGAGCTCTACTTTCTG  
TTTTTTACTTTTCATGATTGGCTGTCTTCCATTTATTCTGGTCATTTATTGCTAGTGACACTGT  
GCCTGCTTCCAGTAGTCTCATTTTCCCTATTTTGCTAATTTGTTACTTTTTCTTTGCTAATTTGG  
AAGATTAACTCATTTTTTAATAAAATTATGTCTAAGATTAAAAA  
AAA

## **FIGURE 18**

MRVRIGLTLLLCAVLLSLASASSDEEGSQDES LDSKTTLTSDSVKDHTTAGRVVAGQIFLDSEE  
SELESSIQEEEDSLKSQEGESVTEDISFLESPNPNENKDYEEP KKV RKPALTAIEGTAHGEPCHFP  
FLFLDKEYDECTSDGREDGRLWCATTYDYKADEKWGFCETEEEEAAKRRQM QEAE MMYQTGMKILN  
GSNKKSQKREAYRYLQKAASMNHTKALERSYALLFGDYLPQNIQAAREMF EKLTEEGSPKGQTA  
LGFLYASGLGVNSSQAKALVYYTFGALGGNLI AHMVLVSRL

**Important features:**

**Signal peptide:**

amino acids 1-21

**N-glycosylation sites.**

amino acids 195-199, 217-221, 272-276

**Tyrosine kinase phosphorylation site.**

amino acids 220-228

**N-myristoylation sites.**

amino acids 120-126, 253-259, 268-274, 270-274, 285-291, 289-295

**Glycosaminoglycan attachment site.**

amino acids 267-271

**Microbodies C-terminal targeting signal.**

amino acids 299-303

**Type II fibronectin collagen-binding domain protein.**

amino acids 127-169

**Fructose-bisphosphate aldolase class-II protein.**

amino acids 101-119

**FIGURE 19**

AATTCAGATTTTAAGCCCATTTCTGCAGTGGAATTTTCATGAACTAGCAAGAGGACACCATCTTCTT  
GTATTATACAAGAAAGGAGTGTACCTATCACACACAGGGGGAAAAATGCTCTTTTGGGTGCTAGG  
CCTCCTAATCCTCTGTGGTTTTCTGTGGACTCGTAAAGGAAAATAAGATTGAAGACATCACTG  
ATAAGTACATTTTTATCACTGGATGTGACTCGGGCTTTGGAACTTGGCAGCCAGAACTTTTGAT  
AAAAAGGGATTTTCATGTAATCGCTGCCTGTCTGACTGAATCAGGATCAACAGCTTTAAAGGCAGA  
AACCTCAGAGAGACTTCGTACTGTGCTTCTGGATGTGACCGACCCAGAGAATGTCAAGAGGACTG  
CCCAGTGGGTGAAGAACCAAGTTGGGGAGAAAGGTCTCTGGGGTCTGATCAATAATGCTGGTGTT  
CCCGGCGTGCTGGCTCCCCTGACTGGCTGACACTAGAGGACTACAGAGAACCTATTGAAGTGAA  
CCTGTTTGGACTCATCAGTGTGACACTAAATATGCTTCCTTTGGTCAAGAAAGCTCAAGGGAGAG  
TTATTAATGTCTCCAGTGTTGGAGGTCGCCTTGCAATCGTTGGAGGGGGCTATACTCCATCCAAA  
TATGCAGTGGAAGGTTTCAATGACAGCTTAAGACGGGACATGAAAGCTTTTGGTGTGCACGTCTC  
ATGCATTGAACCAGGATTGTTCAAAACAACTTGGCAGATCCAGTAAAGGTAATTGAAAAAAAC  
TCGCCATTTGGGAGCAGCTGTCTCCAGACATCAAACAACAATATGGAGAAGGTTACATTGAAAAA  
AGTCTAGACAACTGAAAGGCAATAAATCCTATGTGAACATGGACCTCTCTCCGGTGGTAGAGTG  
CATGGACCACGCTCTAACAAGTCTCTCCCTAAGACTCATTATGCCGCTGGAAAAGATGCCAAAA  
TTTTCTGGATACCTCTGTCTCACATGCCAGCAGCTTTGCAAGACTTTTTATTGTTGAAACAGAAA  
GCAGAGCTGGCTAATCCCAAGGCAGTGTGACTCAGCTAACCACAAATGTCTCCTCCAGGCTATGA  
AATTGGCCGATTTCAAGAACACATCTCCTTTTCAACCCATTCTTATCTGCTCCAACCTGGACT  
CATTTAGATCGTGCTTATTTGGATTGCAAAAGGGAGTCCCACCATCGCTGGTGGTATCCCAGGGT  
CCCTGCTCAAGTTTTCTTTGAAAAGGAGGGCTGGAATGGTACATCACATAGGCAAGTCCTGCCCT  
GTATTTAGGCTTTGCCTGCTTGGTGTGATGTAAGGGAAATTGAAAGACTTGCCCATTCAAAATGA  
TCTTTACCGTGGCCTGCCCCATGCTTATGGTCCCCAGCATTTACAGTAACTTGTGAATGTTAAGT  
ATCATCTCTTATCTAAATATTAAAAGATAAGTCAACCCAAAAAAAAAAAAAAAAAAAAAAAAA  
AAAAAAAAAAAAA

## **FIGURE 20**

MLFWVLGLLILCGFLWTRKGKLIKIEDITDKYIFITGCDSGFGNLAARTFDKKGFHVIAACLTESG  
STALKAETSERLRTVLLDVTDPENVKRTAQWVKNQVGEKGLWGLINNAGVPGVLAPTDWLTLEDY  
REPIEVNLFGLISVTNLMLPLVKKAQGRVINVSSVGGRLAIVGGGYTPSKYAVEGFNDSLRRDMK  
AFGVHVSCIEPGLFKTNLADPVKVIEKKLAIWEQLSPDIKQQYGEGYIEKSLDKLKGNKSYVNMD  
LSPVVECMDHALTSLFPKTHYAAGKDAKIFWIPLSHMPAALQDFLLLKQKAELANPKAV

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-17

#### **Transmembrane domain:**

amino acids 136-152

#### **N-glycosylation sites.**

amino acids 161-163, 187-190 and 253-256

#### **Glycosaminoglycan attachment site.**

amino acids 39-42

#### **N-myristoylation sites.**

amino acids 36-41, 42-47, 108-113, 166-171, 198-203 and 207-212

**FIGURE 21**

CTGAGGCGGCGGTAGC**ATG**GAGGGGGAGAGTACGTCGGCGGTGCTCTCGGGCTTTGTGCTCGGCG  
 CACTCGCTTTCCAGCACCTCAACACGGACTCGGACACGGAAGGTTTTCTTCTTGGGGAAGTAAAA  
 GGTGAAGCCAAGAACAGCATTACTGATTCCCAAATGGATGATGTTGAAGTTGTTTATACAATTGA  
 CATTCAGAAATATATTCCATGCTATCAGCTTTTTAGCTTTTATAATTCTTCAGGCGAAGTAAATG  
 AGCAAGCACTGAAGAAAATATTATCAAATGTCAAAAAGAATGTGGTAGGTTGGTACAAATTCCGT  
 CGTCATTTCAGATCAGATCATGACGTTTAGAGAGAGGCTGCTTCACAAAACTTGCAGGAGCATTT  
 TTCAAACCAAGACCTTGTTTTTCTGCTATTAACACCAAGTATAATAACAGAAAGCTGCTCTACTC  
 ATCGACTGGAACATTCCTTATATAAACCTCAAAAAGGACTTTTTTCACAGGGTACCTTTAGTGTT  
 GCCAATCTGGGCATGTCTGAACAACTGGGTTATAAACTGTATCAGGTTCTGTATGTCCACTGG  
 TTTTAGCCGAGCAGTACAAACACACAGCTCTAAATTTTTTTGAAGAAGATGGATCCTTAAAGGAGG  
 TACATAAGATAAAATGAAATGTATGCTTCATTACAAGAGGAATTAAAGAGTATATGCAAAAAAGTG  
 GAAGACAGTGAACAAGCAGTAGATAAACTAGTAAAGGATGTAAACAGATTAAAACGAGAAATTGA  
 GAAAAGGAGAGGAGCACAGATTCAGGCAGCAAGAGAGAAGAACATCCAAAAGACCCTCAGGAGA  
 ACATTTTTCTTTGTCAGGCATTACGGACCTTTTTTCCAAATTCTGAATTTCTTCATTCATGTGTT  
 ATGTCTTTAAAAAATAGACATGTTTCTAAAAGTAGCTGTAACCTACAACCACCATCTCGATGTAGT  
 AGACAATCTGACCTTAATGGTAGAACACACTGACATTCCTGAAGCTAGTCCAGCTAGTACACCAC  
 AAATCATTAAGCATAAAGCCTTAGACTTAGATGACAGATGGCAATTCAGAGATCTCGGTTGTTA  
 GATACACAAGACAAACGATCTAAAGCAAATACTGGTAGTAGTAACCAAGATAAAGCATCCAAAAT  
 GAGCAGCCCAGAAACAGATGAAGAAATTGAAAAGATGAAGGGTTTTGGTGAATATTCACGGTCTC  
 CTACATTT**TGA**TCCTTTTAAACCTTACAAGGAGATTTTTTTATTTGGCTGATGGGTAAAGCCAAAC  
 ATTTCTATTGTTTTTACTATGTTGAGCTACTTGCAGTAAGTTCATTTGTTTTTACTATGTTCCACC  
 TGTTTGCAGTAATACACAGATAACTCTTAGTGCAATTTACTTCACAAAGTACTTTTTTCAAACATCA  
 GATGCTTTTATTTCCAAACCTTTTTTTTACCTTTCACTAAGTTGTTGAGGGGAAGGCTTACACAG  
 ACACATTCTTTAGAATTGGAAAAGTGAGACCAGGCACAGTGGCTCACACCTGTAATCCCAGCACT  
 TAGGGAAGACAAGTCAGGAGGATTGATTGAAGCTAGGAGTTAGAGACCAGCCTGGGCAACGTATT  
 GAGACCATGTCTATTAAAAAATAAAATGGAAAAGCAAGAATAGCCTTATTTTCAAATATGGAAA  
 GAAATTTATATGAAAATTTATCTGAGTCATTAAAATTCTCCTTAAGTGATACTTTTTTAGAAGTA  
 CATTATGGCTAGAGTTGCCAGATAAAATGCTGGATATCATGCAATAAATTTGCAAAACATCATCT  
 AAAATTTAAAAAAAAAAAAAAAAAAAAAAAAA

## **FIGURE 22**

MEGESTSAVLSGFVLGALAFQHLNTDSDTEGFLLGEVKGEAKNSITDSQMDDVEVVYTIDIQKYI  
PCYQLFSFYNSSGEVNEQALKKILSNVKKNVVGWYKFRRHSDQIMTFRERLLHKNLQEHFSNQDL  
VFLLLTPSIITESCSTHRLEHSLYKPQKGLFHRVPLVVANLGMSEQLGYKTVSGSCMSTGFSRAV  
QTHSSKFFEEEDGSLKEVHKINEMYASLQEELKSICKKVEDSEQAVDKLVKDVNRLKREIEKRRGA  
QIQAAREKNIQKDPQENIFLCQALRTFFPNSEFLHSCVMSLKNRHVSKSSCNYNHHLDVVDNLT  
MVEHTDIPEASPASTPQIIKHKALDDDRWQFKRSRLDQTQDKRSKANTGSSNQDKASKMSSPET  
DEEIEKMKGFGEYSRSPTE

**Important features:**

**Signal peptide:**

amino acids 1-19

**N-glycosylation sites.**

amino acids 75-79, 322-326

**N-myristoylation site.**

amino acids 184-154

**Growth factor and cytokines receptors family.**

amino acids 134-150



**FIGURE 23**

GGCACAGCCGCGCGGGCGGAGGGCAGAGTCAGCCGAGCCGAGTCCAGCCGGACGAGCGGACCAGCGCAGGGCAGCCCAA  
GCAGCGCGCAGCGAACGCCCCGCGCCGCCCCACACCCTCTGCGGTCCCCGCGGCGCCTGCCACCCTTCCCTCCTTCCCC  
GCGTCCCCGCGCTCGCCGGCCAGTCAGCTTGCCGGGTTCGCTGCCCCGCGAAACCCGAGGTACCAGCCCGCGCCTCT  
GCTTCCCTGGGCGCGCGCCGCTCCACGCCCTCCTTCTCCCTGGCCCGGCGCCTGGCACCAGGGGACCGTTGCCTGA  
CGCGAGGCCCAGCTCTACTTTTCGCCCCGCGTCTCCTCCGCTGCTCGCCTCTTCCACCAACTCCAACCTCCTTCTCCC  
TCCAGCTCCACTCGCTAGTCCCCGACTCCGCCAGCCCTCGGCCGCTGCCGTAGCGCCGCTTCCCGTCCGGTCCCCAA  
GGTGGGAACGCGTCCGCCCCGGCCCGCACCATGGACGGTTCGGCTTGCCCGCGCTTCTCTGCACCCTGGCAGTGCTC  
AGCGCCGCGCTGCTGGGTGCCGAGCTCAAGTCGAAAAGTTGCTCGGAAGTGCGACGTCTTTACGTGTCCAAAGGCTTC  
AACAAGAACGATGCCCCCTCCACGAGATCAACGGTGATCATTTGAAGATCTGTCCCCAGGGTTCTACCTGCTGCTCT  
CAAGAGATGGAGGAGAAGTACAGCCTGCAAAGTAAAGATGATTTCAAAGTGTGGTCAGCGAACAGTGCAATCATTG  
CAAGCTGTCTTTGCTTACGTTACAAGAAGTTTGATGAATTCCTCAAAGAACTACTTGAAAATGCAGAGAAATCCCTG  
AATGATATGTTTGTGAAGACATATGGCCATTTATACATGCAAATTTCTGAGCTATTTAAAGATCTCTTCGTAGAGTTG  
AAACGTTACTACGTGGTGGGAAATGTGAACCTGGAAGAAATGCTAAATGACTTCTGGGCTCGCCTCCTGGAGCGGATG  
TTCCGCTGGTGAACCTCCAGTACCACTTTACAGATGAGTATCTGGAATGTGTGAGCAAGTATACGGAGCAGCTGAAG  
CCCTTCGGAGATGTCCCTCGCAAATTGAAGCTCCAGGTTACTCGTGCTTTTGTAGCAGCCCGTACTTTTCGCTCAAGGC  
TTAGCGGTTGCGGGAGATGTGCTGAGCAAGGTCTCCGTGGTAAACCCACAGCCCAGTGATCCCATGCCCTGTTGAAG  
ATGATCTACTGCTCCCACTGCCGGGGTCTCGTGACTGTGAAGCCATGTTACAACACTACTGCTCAAACATCATGAGAGGC  
TGTTTGGCCAACCAAGGGGATCTCGATTTTGAATGGAACAATTTTCATAGATGCTATGCTGATGGTGGCAGAGAGGCTA  
GAGGGTCCTTTCAACATTGAATCGGTTCATGGATCCCATCGATGTGAAGATTTCTGATGCTATTATGAACATGCAGGAT  
AATAGTGTTCAGTGTCTCAGAAGGTTTTCCAGGGATGTGGACCCCCCAAGCCCCTCCAGCTGGACGAATTTCTCGT  
TCCATCTCTGAAAGTGCCTTCAGTGCTCGCTTCAGACCACATCACCCGAGGAACGCCCAACACAGCAGCTGGCACT  
AGTTTGGACCGACTGGTTACTGATGTCAAGGAGAACTGAAACAGGCCAAGAAATTCTGGTCCTCCCTTCCGAGCAAC  
GTTTGCAACGATGAGAGGATGGCTGCAGGAAACGGCAATGAGGATGACTGTTGGAATGGGAAAGGCAAAAGCAGGTAC  
CTGTTTGCACTGACAGGAAATGGATTAGCCAACCAGGGCAACAACCCAGAGGTCCAGGTTGACACCAGCAAACAGAC  
ATACTGATCCTTCGTCAAATCATGGCTCTTCGAGTGATGACCAGCAAGATGAAGAATGCATACAATGGGAACGACGTG  
GACTTCTTTGATATCAGTGATGAAAGTAGTGGAGAAGGAAGTGAAGTGGCTGTGAGTATCAGCAGTGCCCTTCAGAG  
TTTGACTACAATGCCACTGACCATGCTGGGAAGAGTGCCAATGAGAAAGCCGACAGTGCTGGTGTCCGTCTCGGGCA  
CAGGCCTACCTCCTCACTGTCTTCTGCATCTTGTTCTGGTTATGCAGAGAGAGTGGAGATTAATTCTCAAACCTCTGAG  
AAAAAGTGTTTCATCAAAAAGTTAAAAGGCACCAGTTATCACTTTTCTACCATCCTAGTGACTTTTGCTTTTAAATGAA  
TGGACAACAATGTACAGTTTTTACTATGTGGCCACTGGTTTAAAGAGTGCTGACTTTGTTTTCTCATTTCAGTTTTGGG  
AGGAAAAGGGACTGTGCATTGAGTTGGTTCTGCTCCCCCAAACCATGTTAAACGTGGCTAACAGTGTAGGTACAGAA  
CTATAGTTAGTTGTGCATTTGTGATTTTATCACTCTATTATTTGTTTGTATGTTTTTTCTCATTTCGTTTGTGGGTT  
TTTTTTTCCAACGTGTGATCTCGCCTTGTTTCTTACAAGCAAACAGGGTCCCTTCTTGGCACGTAACATGTACGTATT  
TCTGAAATATTAAATAGCTGTACAGAAGCAGGTTTTATTTATCATGTTATCTTATTAAAAGAAAAAGCCCCAAAAGC

## **FIGURE 24**

MARFGLPALLCTLAVLSAALLAAELKSKSCSEVRRRLYVSKGFNKNDAPLHEINGDHLKICPQGST  
CCSQEMEKEYSLQSKDDFKSVVSEQCNHLQAVFASRYKKFDEFFKELLENAEKSLNDMFVKTYGH  
LYMQNSELFKDLFVELKRYVVGNNLEEMLNDFWARLLERMFLVNSQYHFTDEYLECVSKYTE  
QLKPFQDVPRKLKLQVTRAFVAARTFAQGLAVAGDVVSKVSVVNPTAQCTHALLKMIYCSHCRGL  
VTVKPCYNYCSNIMRGCLANQGDLD FEWNNFIDAMLMVAERLEGPFNIESVMDPIDVKISDAIMN  
MQDNSVQVSQKVFQGC GPPKPLPAGRISR SISESAFSARFRPHHPEERPTTAAGTSLDRLVTDVK  
EKLKQAKKFWSSSLPSNVCNDERMAAGNGNEDDCWNGKGSRYLFAVTGNGLANQGNNPEVQVDT  
KPDILILRQIMALRVMTSKMKNAYNGNDVDFDISDESSGEGSGSGCEYQQCPSEFDYNATDHAG  
KSANEKADSAGVRPGAQAYLLTVFCILFLVMQREWR

**Important features:**

**Signal peptide:**

amino acids 1-22

**ATP/GTP-binding site motif A (P-loop).**

amino acids 515-524

**N-glycosylation site.**

amino acids 514-518

**Glycosaminoglycan attachment sites.**

amino acids 494-498, 498-502

**N-myristoylation sites.**

amino acids 63-69, 224-230, 276-282, 438-444, 497-503, 531-537

**Glypicans proteins.**

amino acids 54-75, 105-157, 238-280, 309-346, 423-460, 468-506

**FIGURE 25**

CTCGCCCTCAAATGGGAACGCTGGCCTGGGACTAAAGCATAGACCACCAGGCTGAGTATCCTGAC  
CTGAGTCATCCCCAGGGATCAGGAGCCTCCAGCAGGGAACCTTCCATTATATTCTTCAAGCAACT  
TACAGCTGCACCGACAGTTGCGATGAAAGTTCTAATCTCTTCCCTCCTCCTGTTGCTGCCACTAA  
TGCTGATGTCCATGGTCTCTAGCAGCCTGAATCCAGGGGTCGCCAGAGGCCACAGGGACCGAGGC  
CAGGCTTCTAGGAGATGGCTCCAGGAAGGCGGCCAAGAATGTGAGTGCAAAGATTGGTTCCTGAG  
AGCCCCGAGAAGAAAATTCATGACAGTGTCTGGGCTGCCAAAGAAGCAGTGCCCCTGTGATCATT  
TCAAGGGCAATGTGAAGAAAACAAGACACCAAAGGCACCACAGAAAGCCAAACAAGCATTCCAGA  
GCCTGCCAGCAATTTCTCAAACAATGTCAGCTAAGAAGCTTTGCTCTGCCTTTGTAGGAGCTCTG  
AGCGCCCACTCTTCCAATTAAACATTCTCAGCCAAGAAGACAGTGAGCACACCTACCAGACACTC  
TTCTTCTCCCACTCACTCTCCCACTGTACCCACCCCTAAATCATTCCAGTGCTCTCAAAAAGCA  
TGTTTTTCAAGATCATTTTGTTTGTTGCTCTCTCTAGTGTCTTCTTCTCTCGTCAGTCTTAGCCT  
GTGCCCTCCCCTTACCCAGGCTTAGGCTTAATTACCTGAAAGATTCCAGGAACTGTAGCTTCCT  
AGCTAGTGTCATTTAACCTTAAATGCAATCAGGAAAGTAGCAAACAGAAGTCAATAAATATTTTT  
AAATGTCAAAAAAAAAAAAAAAAAAAAA

## **FIGURE 26**

MKVLISLLLLLPLMLMSMVSSSLNPGVARGHRDRGQASRRWLQEGGQECECKDWFLRAPRRKFM  
TVSGLPKKQCPDHFKGNVKKTRHQRHHRKPNKHSRACQQFLKQCQLRSFALPL

**Important features:**

**Signal peptide:**

amino acids 1-22

**N-myristoylation sites.**

amino acids 27-33, 46-52

**FIGURE 27**

GGACGCCAGCGCCTGCAGAGGCTGAGCAGGGAAAAAGCCAGTGCCCCAGCGGAAGCACAGCTCAG  
AGCTGGTCTGCCATGGACATCCTGGTCCCCTCCTGCAGCTGCTGGTGCTGCTTCTTACCCTGCC  
CCTGCACCTCATGGCTCTGCTGGGCTGCTGGCAGCCCCCTGTGCAAAGCTACTTCCCCTACCTGA  
TGGCCGTGCTGACTCCCAAGAGCAACCGCAAGATGGAGAGCAAGAAACGGGAGCTCTTCAGCCAG  
ATAAAGGGGCTTACAGGAGCCTCCGGGAAAGTGGCCCTACTGGAGCTGGGCTGCGGAACCGGAGC  
CAACTTTTCAATTCTACCCACCGGGCTGCAGGGTCACCTGCCTAGACCCAAATCCCCACTTTGAGA  
AGTTCCTGACAAAGAGCATGGCTGAGAACAGGCACCTCCAATATGAGCGGTTTGTGGTGGCTCCT  
GGAGAGGACATGAGACAGCTGGCTGATGGCTCCATGGATGTGGTGGTCTGCACTCTGGTGCTGTG  
CTCTGTGCAGAGCCCAAGGAAGGTCCTGCAGGAGGTCCGGAGAGTACTGAGACCGGGAGGTGTGC  
TCTTTTTCTGGGAGCATGTGGCAGAACCATATGGAAGCTGGGCCTTCATGTGGCAGCAAGTTTTC  
GAGCCACCTGGAAACACATTGGGGATGGCTGCTGCCTCACCAGAGAGACCTGGAAGGATCTTGA  
GAACGCCCAGTTCTCCGAAATCCAAATGGAACGACAGCCCCCTCCCTTGAAGTGGCTACCTGTTG  
GGCCCCACATCATGGGAAAGGCTGTCAAACAATCTTTCCCAAGCTCCAAGGCACTCATTTGCTCC  
TTCCCCAGCCTCCAATTAGAACAAGCCACCCACCAGCCTATCTATCTTCCACTGAGAGGGACCTTA  
GCAGAATGAGAGAAGACATTTCATGTACCACCTACTAGTCCCTCTCTCCCCAACCTCTGCCAGGGC  
AATCTCTAACTTCAATCCCGCCTTCGACAGTGAAAAAGCTCTACTTCTACGCTGACCCAGGGAGG  
AAACACTAGGACCCTGTTGTATCCTCAACTGCAAGTTTCTGGACTAGTCTCCCAACGTTTGCCTC  
CCAATGTTGTCCCTTTTCCTTCGTTCCTATGGTAAAGCTCCTCTCGCTTTCCTCCTGAGGCTACAC  
CCATGCGTCTCTAGGAACTGGTCACAAAAGTCATGGTGCCTGCATCCCTGCCAAGCCCCCTGAC  
CCTCTCTCCCCACTACCACCTTCTTCCTGAGCTGGGGGCACCAGGGAGAATCAGAGATGCTGGGG  
ATGCCAGAGCAAGACTCAAAGAGGCAGAGGTTTTTGTCTCAAATATTTTTTAATAAATAGACGAA  
ACCACG

## **FIGURE 28**

MDILVPLLQLLVLLLTLPPLHLMALLGCWQPLCKSYFPYLMAVLTPKSNRKMESKKRELFSQIKGL  
TGASGKVALLELGCGTGANFQFYPPGCRVTCCLDPNPHFEKFLTKSMAENRHLQYERFVVPAGEDM  
RQLADGSMDEVVCTLVLCVQSPRKVLQEVRRLRPGGVLEFFWEHVAEPYGSWAFMWQQVFEPW  
KHIGDGCCLTRETWKDLENAQFSEIQMERQPPPLKWLPVGPHIMGKAVKQSFPSKALICSFPSL  
QLEQATHQPIYLPPLRGT

**Important features:**

**Signal peptide:**

amino acids 1-23

**Leucine zipper pattern.**

amino acids 10-32

**N-myristoylation sites.**

amino acids 64-70, 78-84, 80-86, 91-97, 201-207

**FIGURE 29**

CAATGTTTGCCTATCCACCTCCCCAAGCCCCTTTACCT**ATG**CTGCTGCTAACGCTGCTGCTGCT  
GCTGCTGCTGCTTAAAGGCTCATGCTTGGAGTGGGGACTGGTCGGTGCCCAGAAAGTCTCTTCTG  
CCTGACGCCCCCATCAGGGATTGGGCCTTCTTTCCCCCTTCCTTTCTGTGTCTCCTGCCTCAT  
CGGCCTGCCATGACCTGCAGCCAAGCCCAGCCCCGTGGGGAAGGGGAGAAAGTGGGGGATGGC**TA**  
**AG**AAAGCTGGGAGATAGGGAACAGAAGAGGGTAGTGGGTGGGCTAGGGGGGCTGCCTTATTAAA  
GTGGTTGTTTATGATTCTTATACTAATTTATACAAAGATATTAAGGCCCTGTTCAATTAAGAAATT  
GTTCCCTTCCCCTGTGTTCAATGTTTGTAAGATTGTTCTGTGTAAATATGTCTTTATAATAAAC  
AGTTAAAAGCTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

## **FIGURE 30**

MLLLTLLLLLLLLLLKGSCLEWGLVGAQKVSSATDAPIRDWAFFPPSFLCLLPHRPAMTCSQAQPRG  
EGEKVGDG

**Important features:**

**Signal peptide:**

amino acids 1-15

**Growth factor and cytokines receptors family:**

amino acids 3-18



**FIGURE 31**

GTTTGAATTCCTTCAACTATAACCCACAGTCCAAAAGCAGACTCACTGTGTCCCAGGCTACCAAGTT  
CCTCCAAGCAAGTCATTTCCCTTATTTAACCGATGTGTCCCTCAAACACCTGAGTGCTACTCCCT  
ATTTGCATCTGTTTTGATAAATGATGTTGACACCCTCCACCGAATTCTAAGTGGAATC**ATG**TCGG  
GAAGAGATACAATCCTTGGCCTGTGTATCCTCGCATTAGCCTTGTCTTTGGCCATGATGTTTACC  
TTCAGATTCATCACCACCCTTCTGGTTTACATTTTCATTTTATTGGTTATTTTGGGATTGTTGTT  
TGTCTGCGGTGTTTTATGGTGGCTGTATTATGACTATAACCAACGACCTCAGCATAGAATTGGACA  
CAGAAAGGGAAAATATGAAGTGCGTGCTGGGGTTTGCTATCGTATCCACAGGCATCACGGCAGTG  
CTGCTCGTCTTGATTTTTGTTCTCAGAAAGAGAATAAAATTGACAGTTGAGCTTTTCCAAATCAC  
AAATAAAGCCATCAGCAGTGCTCCCTTCCTGCTGTTCCAGCCACTGTGGACATTTGCCATCCTCA  
TTTTCTTCTGGGTCCTCTGGGTGGCTGTGCTGCTGAGCCTGGGAACTGCAGGAGCTGCCCAGGTT  
ATGGAAGGCGGCCAAGTGGAATATAAGCCCCCTTTCGGGCATTTCGGTACATGTGGTTCGTACCATTT  
AATTGGCCTCATCTGGACTAGTGAATTCATCCTTGCGTGCCAGCAAATGACTATAGCTGGGGCAG  
TGGTTACTTGTTATTTCAACAGAAGTAAAAATGATCCTCCTGATCATCCCATCCTTTTCGTCTCTC  
TCCATTCTCTTCTTCTACCATCAAGGAACCGTTGTGAAAGGGTCATTTTTAATCTCTGTGGTGAG  
GATTCCGAGAATCATTGTCATGTACATGCAAAACGCACTGAAAGAACAGCAGCATGGTGCATTGT  
CCAGGTACCTGTTCCGATGCTGCTACTGCTGTTTCTGGTGTCTTGACAAATACCTGCTCCATCTC  
AACCAGAATGCATATACTACAACCTGCTATTAATGGGACAGATTTCTGTACATCAGCAAAAGATGC  
ATTCAAAATCCTTGCCAAGAAGTCAAGTCACTTTACATCTATTAAGTCTTTGGAGACTTCATAA  
TTTTTCTAGGAAAGGTGTTAGTGGTGTGTTTCACTGTTTTTGGAGGACTCATGGCTTTTAACTAC  
AATCGGGCATTCCAGGTGTGGGCAGTCCCTCTGTTATTGGTAGCTTTTTTTGCCTACTTAGTAGC  
CCATAGTTTTTTATCTGTGTTTGAAACTGTGCTGGATGCACTTTTCCTGTGTTTTGCTGTTGATC  
TGGAACAAATGATGGATCGTCAGAAAAGCCCTACTTTATGGATCAAGAATTTCTGAGTTTCGTA  
AAAAGGAGCAACAAATTAAACAATGCAAGGGCACAGCAGGACAAGCACTCATTAAGGAATGAGGA  
GGGAACAGAACTCCAGGCCATTGTGAGAT**TAG**ATACCCATTTAGGTATCTGTACCTGGAAAACATT  
TCCTTCTAAGAGCCATTTACAGAATAGAAGATGAGACCACTAGAGAAAAGTTAGTGAATTTTTTT  
TTAAAGACCTAATAAACCTATTCTTCCTCAAAA

## **FIGURE 32**

MSGRDTILGLCILALALSLAMMFTFRFITLLVHIFISLVILGLLFVCGVLWWLYDYTNDLSIE  
LDTERENMKCVLGFAIVSTGITAVLLVLIFVLRKRIKLTVELFQITNKAISSAPFLLFQPLWTFA  
ILIFFWVLWVAVLLSLGTAGAAQVMEGGQVEYKPLSGIRYMWSYHLIGLIWTSEFILACQQMTIA  
GAVVTCYFNRSKNDPPDHPILSSLSILFFYHQGTVVKGSFLISVVRIPRIIVMYMQNALKEQQHG  
ALSRYLFRCCYCCFWCLDKYLLHLNQNAITTTAINGTDFCTSAKDAFKILSKNSSHFTSINCFGD  
FIIFLGKVLVVCFTVFGGLMAFNYNRAFQVWAVPLLLVAFFAYLVAHSFLSVFETVLDALFLCFA  
VDLETNDGSSEKPYFMDQEFLSFVKRSNKLNNARAQQDKHSLRNEEGTELQAIVR

**Important features:**

**Signal peptide:**

amino acids 1-20

**Putative transmembrane domains:**

amino acids 35-54, 75-97, 126-146, 185-204, 333-350, 352-371

**N-glycosylation sites.**

amino acids 204-208, 295-299, 313-317

**N-myristoylation sites.**

amino acids 147-153, 178-184, 196-202, 296-275, 342-348

[illegible]

**FIGURE 34**

MRTVVLTMKASVIEMFLVLLVTGVHSNKETAKKIKRPFKFTVPQINCDVKAGKIIDPEFIVKCPAG  
CQDPKYHVGTDVYASYSSVCGAAVHSGVLDNSGGKILVRKVAGQSGYKGSYSNGVQSLSLPRWR  
ESFIVLESKPKKGVITYPSALTYSSSKSPAAQAGETTKAYQRPPIPGTTAQPVTLMQLLAVTVAVA  
TPTTLPRPSPSAASTTSIPRPQSVGHRSEQEMDLWSTATYTSSQNRPRADPGIQRQDPSGAAFQKP  
VGADVSLGLVPKEELSTQSLEPVSLGDPNCKIDLSFLIDGSTSIGKRRFRIQKQLLADVAQALDI  
GPAGPLMGVVQYGDNPATHFNLKTHNTSRDLKTAIEKITQRGGLSNVGRAISFVTKNFFSKANGN  
RSGAPNVVVVMVDGWPTDKVEEASRLARESGINIFFITIEGAAENEKQYVVEPNFANKAVCRTNG  
FYSLHVQSWFGLHKTLLQPLVKRVCDTDRLACSKTCLNSADIGFVIDGSSSVGTGNFRTVLQFVTN  
LTKEFEISDTRIGAVQYTYEQRLFEFGFDKYSSKPDILNAIKRVGYWSGGTSTGAAINFALEQL  
FKKSKPNKRKLMILITDGRSYDDVRIPAMAAHLKGVITYAIGVAWAAQEELEVIATHPARDHSEF  
VDEFNLHQYVPRIIQNICTEFNSQPRN

**Important features:****Signal peptide:**

amino acids 1-26

**Transmembrane domain:**

amino acids 181-200

**N-glycosylation sites.**

amino acids 390-394, 520-524

**N-myristoylation sites.**

amino acids 23-29, 93-99, 115-121, 262-268, 367-373, 389-395,  
431-437, 466-472, 509-515, 570-576, 571-577, 575-581, 627-633

**Amidation site.**

amino acids 304-308

**FIGURE 35**

CCGAGCACAGGAGATTGCCTGCGTTTAGGAGGTGGCTGCGTTGTGGGAAAAGCTATCAAGGAAGAAATTGC  
CAAACCATGTCTTTTTTTCTGTTTTTCAGAGTAGTTTCAACAGATCTGAGTGTTTTAATTAAGCATGGAAT  
ACAGAAAACAACAAAAAACTTAAGCTTTAATTTTCATCTGGAATTCACAGTTTTCTTAGCTCCCTGGACCC  
GGTTGACCTGTTGGCTCTTCCCGCTGGCTGCTCTATCACGTGGTGCTCTCCGACTACTCACCCCGAGTGTA  
AAGAACCTTCGGCTCGCGTGCTTCTGAGCTGCTGTGGATGGCCTCGGCTCTCTGGACTGTCCTTCCGAGTA  
GGATGTCACCTGAGATCCCTCAAATGGAGCCTCCTGCTGCTGCTCACTCCTGAGTTTCTTTGTGATGTGGTAC  
CTCAGCCTTCCCCACTACAATGTGATAGAACGCGTGAAGTGGATGTACTTCTATGAGTATGAGCCGATTTA  
CAGACAAGACTTTCACCTTCACTTCGAGAGCATTCAAAGTCTCTCATCAAAATCCATTTCTGGTCATTC  
TGGTGACCTCCCACCTTCAGATGTGAAAGCCAGGCAGGCCATTAGAGTTACTTGGGGTGAAAAAAGTCT  
TGGTGGGGATATGAGGTTCTTACATTTTTCTTATTAGGCCAAGAGGCTGAAAAGGAAGACAAAATGTTGGC  
ATTGTCCTTAGAGGATGAACACCTTCTTTATGGTGACATAATCCGACAAGATTTTTTTAGACACATATAATA  
ACCTGACCTTGAAAACCATTTATGGCATTTCAGGTGGGTAAGTGGAGTTTGGCCCAATGCCAAGTACGTAATG  
AAGACAGACACTGATGTTTTTCATCAATACTGGCAATTTAGTGAAGTATCTTTTAAACCTAAACCACTCAGA  
GAAGTTTTTTCACAGGTTATCCTCTAATTGATAATTATTCCTATAGAGGATTTTACCAAAAAACCCATATTT  
CTTACCAGGAGTATCCTTTCAAGGTGTTCCCTCCATACTGCAGTGGGTTGGGTTATATAATGTCCAGAGAT  
TTGGTGCCAAGGATCTATGAAATGATGGGTACGTAAAACCCATCAAGTTTGAAGATGTTTATGTCCGGGAT  
CTGTTTGAATTTATTAAAAGTGAACATTCATATTCAGAAAGACACAAATCTTTTCTTTCTATATAGAATCC  
ATTTGGATGTCTGTCAACTGAGACGTGTGATTGCAGCCCATGGCTTTTCTTCCAAGGAGATCATCACTTTT  
TGCGAGGTCATGCTAAGGAACACCACATGCCATTATTAACTTCACATTCTACAAAAAGCCTAGAAGGACAG  
GATACCTTGTGGAAAGTGTTAAATAAAGTAGGTACTGTGGAAAATTCATGGGGAGGTCAGTGTGCTGGCTT  
ACACTGAACTGAACTCATGAAAAACCCAGACTGGAGACTGGAGGGTTACACTTGTGATTTATTAGTCAGG  
CCCTTCAAAGATGATATGTGGAGGAATTAATATAAAGGAATTGGAGGTTTTTGCTAAAGAAATTAATAGG  
ACCAAACAATTTGGACATGTCAATTCGTAGACTAGAATTTCTTAAAGGGTGTTACTGAGTTATAAGCTCA  
CTAGGCTGTAAAAACAAAACAATGTAGAGTTTTATTATTGAACAATGTAGTCACTTGAAGGTTTTGTGTA  
TATCTTATGTGGATTACCAATTTAAAAATATATGTAGTTCTGTGTCAAAAAACTTCTTCACTGAAGTTATA  
CTGAACAAAATTTTACCTGTTTTTGGTCATTTATAAAGTACTTCAAGATGTTGCAGTATTTACAGTTATT  
ATTATTTAAATTAATTTCACTTTGTGTTTTTAAATGTTTTGACGATTTCAATACAAGATAAAAAGGATAG  
TGAATCATTTCTTTACATGCAAACATTTTCCAGTTACTTAACTGATCAGTTTATTATTGATACATCACTCCA  
TTAATGTAAAGTCATAGGTCATTATTGCATATCAGTAATCTCTTGGACTTTGTAAATATTTTACTGTGGT  
AATATAGAGAAGAATTAAAGCAAGAAAATCTGAAAA

## **FIGURE 36**

MASALWTVLPSRMSLRSLKWSLLLLSLLSFFVMWYLSLPHYNVIERVNWMYFYEYEPYRQDFHF  
TLREHSNCSHQNPFLVILVTSHPSDVKARQAIRVTWGEKKSWWGYEVLTFLLGQEA EKEDKMLA  
LSLEDEHLLYGDIIRQDFLD TYNNLT LKTIMAFRWVTEFCPNAKYVMKTD TDVFINTGNLVKYLL  
NLNHSEKFFTGYPLIDNYSYRGFYQKTHISYQEYPFKVFPPYCSGLGYIMSRDLVPRIYEMMGHV  
KPIKFEDVYVGICLNLLKVNIHIPEDTNLFFLYRIHLDVCQLRRVIAAHGFSSKEIITFWQVMLR  
NTTCHY

**Important features:**

**Type II transmembrane domain:**

amino acids 20-39

**N-glycosylation sites.**

amino acids 72-76, 154-158, 198-202, 212-216, 326-330

**Glycosaminoglycan attachment site.**

amino acids 239-243

**Ly-6 / u-PAR domain proteins.**

amino acids 23-37

**N-myristoylation site.**

amino acids 271-277

**FIGURE 37**

CGCTCGGGCACCAGCCGCGGCAAGGATGGAGCTGGGTTGCTGGACGCAGTTGGGGCTCACTTTTCTTCAGCTCCTTCT  
CATCTCGTCCTTGCCAAGAGAGTACACAGTCATTAATGAAGCCTGCCCTGGAGCAGAGTGGAATATCATGTGTCTGGGA  
GTGCTGTGAATATGATCAGATTGAGTGCCTCTGCCCCGAAAGAGGGAAGTCGTGGGTATACCATCCCTTGCTGCAG  
GAATGAGGAGAATGAGTGTGACTCCTGCCTGATCCACCCAGGTTGTACCATCTTTGAAAAGCTGCAAGAGCTGCCGAAA  
TGGCTCATGGGGGGGTACCTTGATGACTTCTATGTGAAGGGGTTCTACTGTGCAGAGTGCCGAGCAGGCTGGTACGG  
AGGAGACTGCATGCGATGTGGCCAGGTTCTGCGAGCCCCAAAGGGTCAGATTTTGTGGAAAGCTATCCCCTAAATGC  
TCACTGTGAATGGACCATTTCATGCTAAACCTGGGTTTGTTCATCCAACCTAAGATTTGTTCATGTTGAGTCTGGAGTTTGA  
CTACATGTGCCAGTATGACTATGTTGAGGTTTCGTGATGGAGACAACCGCGATGGCCAGATCATCAAGCGTGTCTGTGG  
CAACGAGCGGCCAGCTCCTATCCAGAGCATAGGATCCTCACTCCACGTCCTCTTCCACTCCGATGGCTCCAAGAATTT  
TGACGGTTTTCATGCCATTTATGAGGAGATCACAGCATGCTCCTCATCCCCTTGTTTTCATGACGGCACGTGCGTCTCT  
TGACAAGGCTGGATCTTACAAGTGTGCCTGCTTGGCAGGCTATACTGGGCAGCGCTGTGAAAATCTCCTTGAAGAAAG  
AAACTGCTCAGACCCTGGGGGGCCAGTCAATGGGTACCAGAAAATAACAGGGGGGCCCTGGGCTTATCAACGGACGCCA  
TGCTAAAATTGGCACCGTGGTGTCTTTCTTTTGTAACTCCTATGTTCTTAGTGGCAATGAGAAAAGAACTTGCCA  
GCAGAAATGGAGAGTGGTCAGGGAAACAGCCCATCTGCATAAAAGCCTGCCGAGAACCAGAAATTTAGACCTGGTGAG  
AAGGAGAGTTCTTCCGATGCAGGTTCAAGGGAGACACCATTAACACAGCTATACTCAGCGGCCCTTCAGCAAGCA  
GAAACTGCAGAGTGCCCCCTACCAAGAAGCCAGCCCTTCCCTTTGGAGATCTGCCCATGGGATACCAACATCTGCATAC  
CCAGCTCCAGTATGAGTGCATCTCACCTTCTACCGCCGCTGGGCAGCAGCAGGAGGACATGTCTGAGGACTGGGAA  
GTGGAGTGGGCGGGCACCATCCTGCATCCCTATCTGCGGGAAAATTGAGAACATCACTGCTCCAAAGACCCAAGGGTT  
GCGCTGGCCGTGGCAGGCAGCCATCTACAGGAGGACCAGCGGGGTGCATGACGGCAGCCTACACAAGGGAGCGTGGTT  
CCTAGTCTGCAGCGGTGCCCTGGTGAATGAGCGCACTGTGGTGGTGGCTGCCCACTGTGTTACTGACCTGGGGAAGGT  
CACCATGATCAAGACAGCAGACCTGAAAGTTGTTTTGGGGAAATTTCTACCGGGATGATGACCGGGATGAGAAGACCAT  
CCAGAGCCTACAGATTTCTGCTATCATTTCTGCATCCCAACTATGACCCCATCCTGCTTGATGCTGACATCGCCATCCT  
GAAGCTCCTAGACAAGGCCCGTATCAGCACCCGAGTCCAGCCCATCTGCCTCGCTGCCAGTCGGGATCTCAGCACTTC  
CTTCCAGGAGTCCCACATCACTGTGGCTGGCTGGAATGTCTGGCAGACGTGAGGAGCCCTGGCTTCAAGAACGACAC  
ACTGCGCTCTGGGGTGGTCACTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT  
TGTCATGATAACATGTTCTGTGCCAGCTGGGAACCCACTGCCCCCTTCTGATATCTGCACTGCAGAGACAGGAGGCAT  
CGCGGCTGTGTCTTCCCGGGACGAGCATCTCCTGAGCCACGCTGGCATCTGATGGGACTGGTCAGCTGGAGCTATGA  
TAAACATGCAGCCACAGGCTCTCCACTGCCTTCACCAAGGTGCTGCCTTTTAAAGACTGGATTGAAAGAAATATGAA  
ATGAACCATGCTCATGCACTCCTTGAGAAGTGTTTTCTGTATATCCGTCTGTACGTGTGTCAATTGCGTGAAGCAGTGTG  
GGCCTGAAGTGTGATTTGGCCTGTGAACTTGGCTGTGCCAGGGCTTCTGACTTCAGGGACAAAAGCTCAGTGAAGGGTG  
AGTAGACCTCCATTGCTGGTAGGCTGATGCCGCGTCCACTACTAGGACAGCCAATTGGAAGATGCCAGGGCTTGCAAG  
AAGTAAGTTTCTTCAAAGAAGACCATATACAAAACCTCTCCACTCCACTGACCTGGTGGTCTTCCCCAACTTTTCAGTT  
ATACGAATGCCATCAGCTTGACCAGGGAAGATCTGGGCTTCATGAGGCCCTTTTGAGGCTCTCAAGTTCTAGAGAGC  
TGCCTGTGGGACAGCCCAGGGCAGCAGAGCTGGGATGTGGTGCATGCCTTTGTGTACATGGCCACAGTACAGTCTGGT  
CCTTTTCTTCCCCATCTCTTGTACACATTTTAATAAAATAAGGGTTGGCTTCTGAACTACAAAAAAAAAAAAAAAAAAAA  
AAA  
AAA

## **FIGURE 38**

MELGCWTQLGLTFLQLLLISSLPREYTVINEACPGAENIMCRECCEYDQIECVCPGKREVVGYT  
 IPCCRNEENECDSCLIHPGCTIFENCKSCRNGSWGGTLDDFYVKGIFYCAECRAGWYGGDCMRCGQ  
 VLRAPKGQILLESYPLNAHCEWTIHAKPGFVIQLRFVMLSLEFDYMCQYDYVEVRDGDNRDQII  
 KRVCGNERPAPIQSIGSSLHVLFSHSDGSKNFDGFHAIYEEITACSSSPCFHDGTCVLDKAGSYKC  
 ACLAGYTGQRCENLLEERNCSDFGGPVNGYQKITGGPGLINGRHAKIGTVVSFFCNNSYVLSGNE  
 KRTCQQNGEWSGKQPICIKACREPKISDLVRRRVLPMQVQSRETPLHQLYSAAFSKQKLQSAPTK  
 KPALPFGDLPMGYQHLHTQLQYECISPFYRRLGSSRRTCLRTGKWSGRAPSCIPICGKIENITAP  
 KTQGLRWPWQAAIYRRTSGVHDGSLHKGAWFLVCSGALVNERTVVVAAHCVTDLGKVTMIKTADL  
 KVVLGKFYRDDDRDEKTIQSLQISAIILHPNYDPILLDADIAILKLLDKARISTRVQPICLAASR  
 DLSTSFQESHITVAGWNVLADVRSFGFKNDTLRSGVVSVDSSLCEEQHEDHGIPVSVTDNMFCA  
 SWEPTAPSDICTAETGGIAAVSFPGRASPEPRWHLMGLVSWSYDKTCSHRLSTAFTKVLFPKDWI  
 ERNMK

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-23

#### **EGF-like domain cysteine pattern signature.**

amino acids 260-272

#### **N-glycosylation sites.**

amino acids 96-100, 279-283, 316-320, 451-455, 614-618

#### **N-myristoylation sites.**

amino acids 35-41, 97-103, 256-262, 284-290, 298-304, 308-314,  
 474-480, 491-497, 638-644, 666-672

#### **Amidation site.**

amino acids 56-60

#### **Serine proteases, trypsin family.**

amino acids 489-506

#### **CUB domain proteins profile.**

amino acids 150-167



**FIGURE 39**

GGTTCCTACATCCTCTCATCTGAGAATCAGAGAGCATAATCTTCTTACGGGCCCCGTGATTTATTAACGTGGCTTAATC  
TGAAGGTTCTCAGTCAAATTTCTTTGTGATCTACTGATTGTGGGGGCATGGCAAGGTTTGCTTAAAGGAGCTTGGCTGG  
TTTGGGCCCTTGTAGCTGACAGAAGGTGGCCAGGGAGAATGCAGCACACTGCTCGGAGAATGAAGGCGCTTCTGTTGC  
TGGTCTTGCCTTGGCTCAGTCCTGCTAACTACATTGACAATGTGGGCAACCTGCACTTCTGTATTGAGAACTCTGTA  
AAGGTGCCTCCCCTACGGCCTGACCAAAGATAGGAAGAGGCGCTCACAAGATGGCTGTCCAGACGGCTGTGCGAGCC  
TCACAGCCACGGCTCCCTCCCCAGAGGTTTCTGCAGCTGCCACCATCTCCTTAATGACAGACGAGCCTGGCCTAGACA  
ACCCTGCCCTACGTGTCTCGGCAGAGGACGGGCAGCCAGCAATCAGCCCAGTGGACTCTGGCCGGAGCAACCGAACTA  
GGGCACGGCCCTTTGAGAGATCCACTATTAGAAGCAGATCATTTAAAAAATAAATCGAGCTTTGAGTGTCTTCGAA  
GGACAAAGAGCGGGAGTGCAGTTGCCAACCATGCCGACCAGGGCAGGGAAAATTCTGAAAACACCACTGCCCCCTGAAG  
TCTTTCCAAGGTTGTACCACCTGATTCCAGATGGTGAAATTAACCAGCATCAAGATCAATCGAGTAGATCCCAGTGAAA  
GCCTCTCTATTAGGCTGGTGGGAGGTAGCGAAACCCCACTGGTCCATATCATTATCCAACACATTTATCGTGATGGGG  
TGATCGCCAGAGACGGCCGGCTACTGCCAGGAGACATCATTTCAAAGGTCAACGGGATGGACATCAGCAATGTCCCTC  
ACAACTACGCTGTGCGTCTCCTGCGGCAGCCCTGCCAGGTGCTGTGGCTGACTGTGATGCGTGAACAGAAGTTCCGCA  
GCAGGAACAATGGACAGGCCCCGGATGCCTACAGACCCCGAGATGACAGCTTTCATGTGATTCTCAACAAAAGTAGCC  
CCGAGGAGCAGCTTGGAATAAACTGGTGCGCAAGGTGGATGAGCCTGGGGTTTTCATCTTCAATGTGCTGGATGGCG  
GTGTGGCATATCGACATGGTCAGCTTGAGGAGAATGACCGTGTGTTAGCCATCAATGGACATGATCTTCGATATGGCA  
GCCCAGAAAGTGCGGCTCATCTGATTGAGGCCAGTGAAAGACGTGTTACCTCGTCGTGTCCCGCCAGGTTCCGGCAGC  
GGAGCCCTGACATCTTTCAGGAAGCCGGCTGGAACAGCAATGGCAGCTGGTCCCCAGGGCCAGGGGAGAGGAGCAACA  
CTCCCAAGCCCCCTCCATCTTACAATTACTTGTGTCATGAGAAGGTGGTAAATATCCAAAAGACCCCGGTGAATCTCTCG  
GCATGACCGTCGCAGGGGGAGCATCACATAGAGAATGGGATTTGCCATCTATGTGTCATCAGTGTGAGCCCGGAGGAG  
TCATAAGCAGAGATGGAAGAATAAAACAGGTGACATTTTGTGTAATGTGGATGGGGTGAAGTACAGAGGTCAGCC  
GGAGTGAGGCAGTGGCATTATTGAAAAGAACATCATCCTCGATAGTACTCAAAGCTTTGGAAGTCAAAGAGTATGAGC  
CCCAGGAAGACTGCAGCAGCCCAGCAGCCCTGGACTCCAACCACAACATGGCCCCACCCAGTGACTGGTCCCCATCCT  
GGGTGATGTGGCTGGAATTACCACGGTGCTTGTATAACTGTAAAGATATTGTATTACGAAGAAACACAGCTGGAAGTC  
TGGGCTTCTGCATTGTAGGAGGTTATGAAGAATACAATGGAAACAAACCTTTTTTTCATCAAATCCATTGTTGAAGGAA  
CACCAGCATAACAATGATGGAAGAATTAGATGTGGTGATATCTTCTTGCTGTCAATGGTAGAAGTACATCAGGAATGA  
TACATGCTTGCTTGGCAAGACTGCTGAAAGAACTTAAAGGAAGAATTACTCTAACTATTGTTTCTTGGCCTGGCACTT  
TTTTATTAGAATCAATGATGGGTGAGAGGAAAACAGAAAAATCACAAATAGGCTAAGAAGTTGAAACACTATATTTATC  
TTGTGAGTTTTTATATTTAAAGAAAGAATACATTGTAAAAATGTCAGGAAAAGTATGATCATCTAATGAAAGCCAGTT  
ACACCTCAGAAAATATGATTCCAAAAAATTTAAACTACTAGTTTTTTTTTTCAGTGTGGAGGATTTCTCATTACTCTAC  
AACATTGTTTATATTTTTTCTATTCAATAAAAAGCCCTAAAACAATAAAATGATTGATTTGTATACCCCACTGAATT  
CAAGCTGATTTAAATTTAAATTTGGTATATGCTGAAGTCTGCCAAGGGTACATTATGGCCATTTTTTAATTTACAGCT  
AAAATATTTTTTAAATGCATTGCTGAGAAACGTTGCTTTCATCAAACAAGAATAAATATTTTTTCAGAAGTTAAA

## **FIGURE 40**

MKALLLLVLPWLSPANYIDNVGNLHFLYSELCKGASHYGLTKDRKRRSQDGCPCDGCASLTATAPS  
PEVSAAATISLMTDEPGLDNPAYVSSAEDGQPAISPVDSGRSNRTRARPFERSTIRSRSEFKKINR  
ALSVLRRTKSGSAVANHADQGRENSENTTAPEVFPRLYHLIPDGEITSIKINRVDPSESLSIRLV  
GGSETPLVHIIIIQHIIYRDGVIARDGRLLPGDIILKVNGMDISNVPHNYAVRLLRQPCQVLWLTVM  
REQKFRSRNNGQAPDAYRPRDDSFHVILNKSSPEEQLGIKLVKRVDEPGVFIFNVLDGGVAYRHG  
QLEENDRVLAINGHDLRYGSPESAHLIQASERRVHLVVSQRQVRQSPDIFQEAGWNSNGSWSPG  
PGRSNTPKPLHPTITCHEKVNIQKDPGESLGMTVAGGASHREWDLPYVISVEPGGVISRDGR  
IKTGDILLNVLDGVELTEVSRSEAVALLKRTSSSIVLKALEVKEYEPQEDCSSPAALDSNHNMAPP  
SDWSPSWVMWLELPRCLYNCKDIVLRRNTAGSLGFCIVGGYEEYNGNKPFFIKSIVEGTPAYNDG  
RIRCGDILLAVNGRSTSGMIHACLARLLKELKGRITLTIVSWPGTFL

**Important features:**

**Signal peptide:**

amino acids 1-15

**N-glycosylation sites.**

amino acids 108-112, 157-161, 289-293, 384-388

**Tyrosine kinase phosphorylation sites.**

amino acids 433-441, 492-500

**N-myristoylation sites.**

amino acids 51-57, 141-147, 233-239, 344-350, 423-429, 447-453,  
467-473, 603-609

**FIGURE 41**

ACCAGGCATTGTATCTTCAGTTGTCATCAAGTTCGCAATCAGATTGGAAAAGCTCAACTTGAAGC  
TTTCTTGCCTGCAGTGAAGCAGAGAGATAGATATTATTACGTAATAAAAAACATGGGCTTCAAC  
CTGACTTTCCACCTTTCCTACAAATTCCGATTACTGTTGCTGTTGACTTTGTGCCTGACAGTGGT  
TGGGTGGGCCACCAGTAACTACTTCGTGGGTGCCATTCAAGAGATTCCCTAAAGCAAAGGAGTTCA  
TGGCTAATTTCCATAAGACCCTCATTTTGGGGAAGGGAAAACTCTGACTAATGAAGCATCCACG  
AAGAAGGTAGAACTTGACAACTGTCCTTCTGTGTCTCCTTACCTCAGAGGCCAGAGCAAGCTCAT  
TTTCAAACCAGATCTCACTTTTGAAGAGGTACAGGCAGAAAATCCCAAAGTGTCCAGAGGCCGGT  
ATCGCCCTCAGGAATGTAAAGCTTTACAGAGGGTCGCCATCCTCGTTCCCCACCGGAACAGAGAG  
AAACACCTGATGTACCTGCTGGAACATCTGCATCCCTTCCTGCAGAGGCAGCAGCTGGATTATGG  
CATCTACGTCATCCACCAGGCTGAAGGTAAAAAGTTTAATCGAGCCAACTCTTGAATGTGGGCT  
ATCTAGAAGCCCTCAAGGAAGAAAATTGGGACTGCTTTATATTCCACGATGTGGACCTGGTACCC  
GAGAATGACTTTAACCTTTACAAGTGTGAGGAGCATCCCAAGCATCTGGTGGTTGGCAGGAACAG  
CACTGGGTACAGGTTACGTTACAGTGGATATTTTGGGGGTGTTACTGCCCTAAGCAGAGAGCAGT  
TTTTCAAGGTGAATGGATTCTCTAACAACACTACTGGGGATGGGGAGGCGAAGACGATGACCTCAGA  
CTCAGGGTTGAGCTCCAAAGAATGAAAATTTCCCGGCCCTGCCTGAAGTGGGTAAATATACAAT  
GGTCTTCCACACTAGAGACAAAGGCAATGAGGTGAACGCAGAACGGATGAAGCTCTTACACCAAG  
TGTCACGAGTCTGGAGAACAGATGGGTTGAGTAGTTGTTCTTATAAATTAGTATCTGTGGAACAC  
AATCCTTTTATATATCAACATCACAGTGGATTTCTGGTTTGGTGCATTGACCCTGGATCTTTTGGTG  
ATGTTTGAAGAAGTGAATCTTTGTTTGCAATAATTTTGGCCTAGAGACTTCAAATAGTAGCACA  
CATTAGAACCTGTTACAGCTCATTGTTGAGCTGAATTTTTCTTTTTGTATTTTCTTAGCAGAG  
CTCCTGGTGATGTAGAGTATAAAACAGTTGTAACAAGACAGCTTTCTTAGTCATTTTGATCATGA  
GGGTAAATATTGTAATATGGATACTTGAAGGACTTTATATAAAAGGATGACTCAAAGGATAAAA  
TGAACGCTATTTGAGGACTCTGGTTGAAGGAGATTTATTTAAATTTGAAGTAATATATTATGGGA  
TAAAAGGCCACAGGAAATAAGACTGCTGAATGTCTGAGAGAACCAGAGTTGTTCTCGTCCAAGGT  
AGAAAGGTACGAAGATACAATACTGTTATTCATTTATCCTGTACAATCATCTGTGAAGTGGTGGT  
GTCAGGTGAGAAGGCGTCCACAAAAGAGGGGAGAAAAGGCGACGAATCAGGACACAGTGAACCTTG  
GGAATGAAGAGGTAGCAGGAGGGTGGAGTGTGGCTGCAAAGGCAGCAGTAGCTGAGCTGGTTGC  
AGGTGCTGATAGCCTTCAGGGGAGGACCTGCCAGGTATGCCTTCAGTGATGCCACCAGAGAA  
TACATTCTCTATTAGTTTTTTAAAGAGTTTTTGTAAAATGATTTTGTACAAGTAGGATATGAATTA  
GCAGTTTACAAGTTTACATATTAATAATAATAATATGTCTATCAAATACCTCTGTAGTAAAAT  
GTGAAAAGCAAAA

## **FIGURE 42**

MGFNLT FHL SYKFRLLLLLTLCLTVVGWATSNYFVGAIQEIPKAKEFMANFHKTLLILGKGKTLTN  
EASTKKVELDNCPSVSPYLRGQSKLIFKPDLTLEEVQAENPKVSRGRYRPQECKALQRVAILVPH  
RNREKHLMYLLEHLHPFLQRQQLDYGIYVIHQAEKKFNRAKLLNVGYLEALKEENWDCFI FHDV  
DLVPENDFNLYKCEEHPKHLVVGRNSTGYRLRYSGYFGGVTALSREQFFKVNGFSNNYWG WGGED  
DDLRLRVELQRMKISRPLPEVGKYTMVFHTRDKGNEVNAERMKLLHQVSRVWRTDGLSSCSYKLV  
SVEHNPLYINITVDFWFGA

**Important features:**

**Signal peptide:**

amino acids 1-27

**N-glycosylation sites.**

amino acids 4-8, 220-224, 335-339

**Xylose isomerase proteins.**

amino acids 191-202

**FIGURE 43**

GCTCAAGACCCAGCAGTGGGACAGCCAGACAGACGGCACG**ATG**GCACTGAGCTCCCAGATCTGGG  
CCGCTTGCCTCCTGCTCCTCCTCCTCCTCGCCAGCCTGACCAGTGGCTCTGTTTTCCCACAACAG  
ACGGGACAACCTGCAGAGCTGCAACCCCAGGACAGAGCTGGAGCCAGGGCCAGCTGGATGCCCAT  
GTTCCAGAGGCGAAGGAGGCGAGACACCCACTTCCCCATCTGCATTTTCTGCTGCGGCTGCTGTC  
ATCGATCAAAGTGTGGGATGTGCTGCAAGACG**TAG**AACCTACCTGCCCTGCCCCGTCCCCTCCC  
TTCCTTATTTATTCCTGCTGCCCCAGAACATAGGTCTTGGAATAAAATGGCTGGTTCTTTTGTTT  
TCCAAA  
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

## **FIGURE 44**

MALSSQIWAACLLLLLLLLASLTSGSVFPQQTGQLAELQPQDRAGARASWMPMFQRRRRRDTHFPI  
CIFCCGCCCHRSKCGMCCKT

**Important features:**

**Signal peptide:**

amino acids 1-24

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 58-59

**N-myristoylation site.**

amino acids 44-50

**Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 1-12

**FIGURE 45**

GTGGCTTCATTTTCAGTGGCTGACTTCCAGAGAGCAATATGGCTGGTTCCCCAACATGCCTCACCC  
TCATCTATATCCTTTGGCAGCTCACAGGGTCAGCAGCCTCTGGACCCGTGAAAGAGCTGGTCGGT  
TCCGTTGGTGGGGCCGTGACTTTCCCCCTGAAGTCCAAAGTAAAGCAAGTTGACTCTATTGTCTG  
GACCTTCAACACAACCCCTCTTGTCACCATAACAGCCAGAAGGGGGCACTATCATAGTGACCCAAA  
ATCGTAATAGGGAGAGAGTAGACTTCCCAGATGGAGGCTACTCCCTGAAGCTCAGCAAACCTGAAG  
AAGAATGACTCAGGGATCTACTATGTGGGGATATACAGCTCATCACTCCAGCAGCCCTCCACCCA  
GGAGTACGTGCTGCATGTCTACGAGCACCTGTCAAAGCCTAAAGTCACCATGGGTCTGCAGAGCA  
ATAAGAATGGCACCTGTGTGACCAATCTGACATGCTGCATGGAACATGGGGAAGAGGATGTGATT  
TATACCTGGAAGGCCCTGGGGCAAGCAGCCAATGAGTCCCATAATGGGTCCATCCTCCCCATCTC  
CTGGAGATGGGGAGAAAGTGATATGACCTTCATCTGCGTTGCCAGGAACCCTGTCAGCAGAACT  
TCTCAAGCCCCATCCTTGCCAGGAAGCTCTGTGAAGGTGCTGCTGATGACCCAGATTCTCCATG  
GTCCTCCTGTGTCTCCTGTTGGTGCCCCCTCTGCTCAGTCTCTTTGTACTGGGGCTATTTCTTTG  
GTTTCTGAAGAGAGAGAGACAAGAAGAGTACATTGAAGAGAAGAAGAGAGTGGACATTTGTCGGG  
AAACTCCTAACATATGCCCCCATCTCTGGAGAGAACACAGAGTACGACACAATCCCTCACACTAAT  
AGAACAATCCTAAAGGAAGATCCAGCAAATACGGTTTACTCCACTGTGGAAATACCGAAAAAGAT  
GGAAAATCCCCACTCACTGCTCACGATGCCAGACACACCAAGGCTATTTGCCTATGAGAATGTTA  
TCTAGACAGCAGTGCACTCCCCTAAGTCTCTGCTCA

## **FIGURE 46**

MAGSPTCLTLIYILWQLTGSAASGPVKELVGSVGGAVTFPLKSKVKQVDSIVWTFNTTPLVTIQP  
EGGTIIIVTQNRNRERVDFPDGGYSLKLSKLKKNDSGIYYVGIYSSSLQQPSTQEYVLHVYEHLSK  
PKVTMGLQSNKNGTCVTNLTCMEHGEEVDIYTWKALGQAANESHNGSILPISWRWGESDMTFIC  
VARNPVSRNFSSPILARKLCEGAADDPDSSMVLLCLLLVPLLLSLFVLGLFLWFLKRERQEEYIE  
EKKRVDICRETPNICPHSGENTEYDTIPHTNRTILKEDPANTVYSTVEIPKKMENPHSLLTMPDT  
PRLFAYENVI

**Important features:**

**Signal peptide:**

amino acids 1-22

**Transmembrane domain:**

amino acids 224-250

**Leucine zipper pattern.**

amino acids 229-251

**N-glycosylation sites.**

amino acids 98-102, 142-146, 148-152, 172-176, 176-180, 204-208,  
291-295



**FIGURE 47**

GGCTCGAGCGTTTCTGAGCCAGGGGTGACCATGACCTGCTGCGAAGGATGGACATCCTGCAATGG  
ATTCAGCCTGCTGGTTCTACTGCTGTAGGAGTAGTTCTCAATGCGATACCTCTAATTGTCAGCT  
TAGTTGAGGAAGACCAATTTTCTCAAACCCCATCTCTTGCTTTGAGTGGTGGTTCCCAGGAATT  
ATAGGAGCAGGTCTGATGGCCATTCCAGCAACAACAATGTCCTTGACAGCAAGAAAAAGAGCGTG  
CTGCAACAACAGAACTGGAATGTTTCTTTCATCATTTTTTCAGTGTGATCACAGTCATTGGTGCTC  
TGTATTGCATGCTGATATCCATCCAGGCTCTCTTAAAAGGTCCTCTCATGTGTAATTCTCCAAGC  
AACAGTAATGCCAATTGTGAATTTTCATTGAAAAACATCAGTGACATTCATCCAGAATCCTTCAA  
CTTGCAAGTGGTTTTTCAATGACTCTTGTGCACCTCCTACTGGTTTTCAATAAACCCACCAGTAACG  
ACACCATGGCGAGTGGCTGGAGAGCATCTAGTTTCCACTTCGATTCTGAAGAAAACAAACATAGG  
CTTATCCACTTCTCAGTATTTTTAGGTCTATTGCTTGTTGGAATTCTGGAGGTCCTGTTTGGGCT  
CAGTCAGATAGTCATCGGTTTCCTTGGCTGTCTGTGTGGAGTCTCTAAGCGAAGAAGTCAAATTG  
TGTAGTTTAATGGGAATAAAATGTAAGTATCAGTAGTTTGAAAAAAAAAAAA

## **FIGURE 48**

MTCCEGWTSCNGFSLLVLLLLGVVLNAIPLIVSLVEEDQFSQNPISCFEWWFPGIIGAGLMAIPA  
TTMSLTARKRACCNNRTGMFLSSFFSVITVIGALYCMLISIQALLKGPLMCNSPSNSNANCEFSL  
KNISDIHPESFNLQWFFNDSCAPPTGFNKPTSNDTMASGWRASSFHFDSEENKHRLIHFSVFLGL  
LLVGILEVLFLGLSQIVIGFLGCLCGVSKRRSQIV

**Important features:**

**Transmembrane domains:**

amino acids 10-31 (type II), 50-72, 87-110, 191-213

**N-glycosylation sites.**

amino acids 80-84, 132-136, 148-152, 163-167

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 223-227

**N-myristoylation sites.**

amino acids 22-28, 54-60, 83-89, 97-103, 216-222

**Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 207-218

**TNFR/NGFR family cysteine-rich region protein.**

amino acids 4-12

**FIGURE 49**

ATCCGTTCTCTGCGCTGCCAGCTCAGGTGAGCCCTCGCCAAGGTGACCTCGCAGGACACTGGTGA  
AGGAGCAGTGAGGAACCTGCAGAGTCACACAGTTGCTGACCAATTGAGCTGTGAGCCTGGAGCAG  
ATCCGTGGGCTGCAGACCCCCGCCCCAGTGCCTCTCCCCCTGCAGCCCTGCCCCCTCGAACTGTGA  
C**ATG**GAGAGAGTGACCCTGGCCCTTCTCCTACTGGCAGGCCTGACTGCCTTGGAAGCCAATGACC  
CATTTGCCAATAAAGACGATCCCTTCTACTATGACTGGAAAAACCTGCAGCTGAGCGGACTGATC  
TGCGGAGGGCTCCTGGCCATTGCTGGGATCGCGGCAGTTCTGAGTGGCAAATGCAAATACAAGAG  
CAGCCAGAAGCAGCACAGTCCTGTACCTGAGAAGGCCATCCCACTCATCACTCCAGGCTCTGCCA  
CTACTTGCT**TGA**GCACAGGACTGGCCTCCAGGGATGGCCTGAAGCCTAACACTGGCCCCCAGCACC  
TCCTCCCCTGGGAGGCCTTATCCTCAAGGAAGGACTTCTCTCCAAGGGCAGGCTGTTAGGCCCCT  
TTCTGATCAGGAGGCTTCTTTATGAATTAACTCGCCCCACCACCCCCTCA

## **FIGURE 50**

MERVTLALLLLAGLTALEANDPFANKDDPFYYDWKNLQLSGLICGGLLAIAGIAAVLSGKCKYKS  
SQKQHSPVPEKAIPPLITPGSATTC

**Important features:**

**Signal peptide:**

amino acids 1-16

**Transmembrane domain:**

amino acids 36-59

**N-myristoylation sites.**

amino acids 41-47, 45-51, 84-90

**Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7.**

amino acids 54-67

## FIGURE 51

[illegible]

## **FIGURE 52**

MKFQGPLACLLLLALCLGSGEAGPLQSGEESTGTNIGEALGHGLGDALSEGVGKAIGKEAGGAAGS  
KVSEALGQGTREAVGTGVRQVPFGAADALGNRVGEAAHALGNTGHEIGRQAEDVIRHGADAVRG  
SWQGVPGHSGAWETSGGHGIFGSQGGLGGQGQGNPGGLGTPWVHGYPGNSAGSFGMNPQGAPWGQ  
GGNGGPPNFGTNTQGAVAQPGYGSVRASNQNEGCTNPPPSGSGGGSSNSGGGSGSQSGSSGSGSN  
GDNNNGSSSGSSSSGSSSSGSSSGSSSGSSSGNSGGSRGDSGSESSWGSSTGSSSGNHGGSG  
GGNGHKPGCEKPGNEARGSGESGIQGFRGQGVSSNMREISKEGNRLLGGSGDNYRGQSSWGS  
GDAVGGVNTVNSETSPGMFNFDTFWKNFKSKLGFINWDAINKDQRSSRIP

**Signal peptide:**

amino acids 1-21

**N-glycosylation site.**

amino acids 265-269

**Glycosaminoglycan attachment site.**

amino acids 235-239, 237-241, 244-248, 255-259, 324-328, 388-392

**Casein kinase II phosphorylation site.**

amino acids 26-30, 109-113, 259-263, 300-304, 304-308

**N-myristoylation site.**

amino acids 17-23, 32-38, 42-48, 50-56, 60-66, 61-67, 64-70, 74-80,  
90-96, 96-102, 130-136, 140-146, 149-155, 152-158, 155-161, 159-165,  
163-169, 178-184, 190-196, 194-200, 199-205, 218-224, 236-242,  
238-244, 239-245, 240-246, 245-251, 246-252, 249-252, 253-259,  
256-262, 266-272, 270-276, 271-277, 275-281, 279-285, 283-289,  
284-290, 287-293, 288-294, 291-297, 292-298, 295-301, 298-304,  
305-311, 311-317, 315-321, 319-325, 322-328, 323-329, 325-331,  
343-349, 354-360, 356-362, 374-380, 381-387, 383-389, 387-393,  
389-395, 395-401

**Cell attachment sequence.**

amino acids 301-304

**FIGURE 53**

GGAGAAGAGGTTGTGTGGGACAAGCTGCTCCCGACAGAAGGATGTCGCTGCTGAGCCTGCCCTGG  
CTGGGCCTCAGACCGGTGGCAATGTCCCCATGGCTACTCCTGCTGCTGGTTGTGGGCTCCTGGCT  
ACTCGCCCGCATCCTGGCTTGGACCTATGCCTTCTATAACAACCTGCCGCCGGCTCCAGTGTTTCC  
CACAGCCCCCAAACGGAACCTGGTTTTTGGGGTCACCTGGGCCTGATCACTCCTACAGAGGAGGGC  
TTGAAGGACTCGACCCAGATGTGCGCCACCTATTTCCAGGGCTTTACGGTATGGCTGGGTCCCAT  
CATCCCCTTCATCGTTTTATGCCACCCTGACACCATCCGGTCTATCACCAATGCCTCAGCTGCCA  
TTGCACCCAAGGATAATCTCTTCATCAGGTTCCCTGAAGCCCTGGCTGGGAGAAGGGATACTGCTG  
AGTGGCGGTGACAAGTGGAGCCGCCACCGTCGGATGCTGACGCCCGCCTTCCATTTCAACATCCT  
GAAGTCCTATATAACGATCTTCAACAAGAGTGCAAACATCATGCTTGACAAGTGGCAGCACCTGG  
CCTCAGAGGGCAGCAGTCGTCTGGACATGTTTGAGCACATCAGCCTCATGACCTTGGACAGTCTA  
CAGAAATGCATCTTCAGCTTTGACAGCCATTGTCAGGAGAGGGCCAGTGAATATATTGCCACCAT  
CTTGGAGCTCAGTGCCCTTGTAGAGAAAAGAAGCCAGCATATCCTCCAGCACATGGACTTTCTGT  
ATTACCTCTCCCATGACGGGGCGGCGCTTCCACAGGGCCTGCCGCCTGGTGCATGACTTCACAGAC  
GCTGTCATCCGGGAGCGGCGTCGCACCCTCCCCACTCAGGGTATTGATGATTTTTTTCAAAGACAA  
AGCCAAGTCCAAGACTTTGGATTTCAATTGATGTGCTTCTGCTGAGCAAGGATGAAGATGGGAAGG  
CATTGTCAGATGAGGATATAAGAGCAGAGGCTGACACCTTCATGTTTGGAGGCCATGACACCACG  
GCCAGTGGCCTCTCCTGGGTCTGTACAACCTTGCGAGGCACCCAGAATACCAGGAGCGCTGCCG  
ACAGGAGGTGCAAGAGCTTCTGAAGGACCGCGATCCTAAAGAGATTGAATGGGACGACCTGGCCC  
AGCTGCCCTTCCTGACCATGTGCGTGAAGGAGAGCCTGAGGTTACATCCCCCAGCTCCCTTCATC  
TCCCGATGCTGCACCCAGGACATTGTTCTCCAGATGGCCGAGTCATCCCCAAAGGCATTACCTG  
CCTCATCGATATTATAGGGGTCCATCACAACCCAACTGTGTGGCCGGATCCTGAGGTCTACGACC  
CCTTCCGCTTTGACCCAGAGAACAGCAAGGGGAGGTCACCTCTGGCTTTTATTCCTTTCTCCGCA  
GGGCCCAGGAACCTGCATCGGGCAGGCGTTCGCCATGGCGGAGATGAAAGTGGTCCTGGCGTTGAT  
GCTGCTGCACTTCCGGTTCCTGCCAGACCACACTGAGCCCCGCAGGAAGCTGGAATTGATCATGC  
GCGCCGAGGGCGGGCTTTGGCTGCGGGTGGAGCCCCCTGAATGTAGGCTTGCACTGACTTTCTGAC  
CCATCCACCTGTTTTTTTTGCAGATTGTCATGAATAAAACGGTGCTGTCAA

## **FIGURE 54**

MSLLSLPWLGLRPVAMSPWLLLLLVVGSWLLARILAWTYAFYNNCRRLQCFPQPPKRNWFWGHLG  
LITPTEEGLKDSTQMSATYSQGFTVWLGPIIPFIVLCHPDTIRSITNASAAIAPKDNLFIRFLKP  
WLGEIGILLSGGDKWSRHRMLTPAFHFNILKSYITIFNKSANIMLDKWQHLASEGSSRLDMFEHI  
SLMTLDSLQKCIFSFDSHCQERPSEYIATILELSALVEKRSQHILQHMDFLYYLSHDGRRFHRAC  
RLVHDFTDVIRERRRTLPTQGIDDFKDKAKSKTLD FIDVLLLSKDEDGKALSDEDIRAEADTF  
MFGGHDTTASGLSWVLYNLARHPEYQERCQEVQELLKDRDPKEIEWDDLAQLPFLTMCVKESLR  
LHPPAPFISRCCTQDIVLPDGRVIPKGITCLIDIIGVHHNPTVWPDPEVYDPFRFDPENSKGRSP  
LAFIPFSAGPRNCIGQAFAMAEMKVVLALMLLHFRFLPDHTEPRRKLELIMRAEGGLWLRVEPLN  
VGLQ

**Important features:**

**Transmembrane domains:**

amino acids 13-32 (type II), 77-102

**Cytochrome P450 cysteine heme-iron ligand signature.**

amino acids 461-471

**N-glycosylation sites.**

amino acids 112-116, 168-172



## FIGURE 55

[illegible]

## **FIGURE 56**

MGPVKQLKRMFEPTRLIATIMVLLCFALTLCSAFWWHNKGLALIFCILQSLALTWYSLSFIPFAR  
DAVKKCFVCLA

**Important features:**

**Signal peptide:**

amino acids 1-33

**Type II fibronectin collagen-binding domain protein.**

amino acids 30-72

[illegible]

## **FIGURE 58**

MLCLCLYVPVIGEAQTEFQYFESKGLPAELKSIFKLSVFIPSQEFSTYRQWKQKIVQAGDKDLTG  
QLDFEEFVHYLQDHEKKLRLVFKILDKKNDGRIDAQEIMQSLRDLGVKISEQQAEEKILKSMDKNG  
TMTIDWNEWRDYHLLHPVENIPEIILYWKHSTIFDVGENLTVPDEFTVEERQTGMWWRHLVAGGG  
AGAVSRTCTAPLDRKVLQMQLVHASRSNNMGIVGGFTQMIREGGARSLWRGNGINVLKIAPESAIAK  
FMAYEQIKRLVGSDQETLRIHERLVAGSLAGAIQSSIIYPMEVLKTRMALRKGTGQYSGMLDCARR  
ILAREGVAAFYKGYVPNMLGIIPYAGIDLAVYETLKNWLQHYAVNSADPGVFVLLACGTMSSTC  
QQLASYPLALVRTRMQAQASIEGAPEVTMSSLFKHILRTEGAFLYRGLAPNFMKVIPAVSISYV  
VYENLKITLGVQSR

**Important features:**

**Signal peptide:**

amino acids 1-16

**Putative transmembrane domains:**

amino acids 284-304, 339-360, 376-394

**Mitochondrial energy transfer proteins signature.**

amino acids 206-215, 300-309

**N-glycosylation sites.**

amino acids 129-133, 169-173

**Elongation Factor-hand calcium-binding protein.**

amino acids 54-73, 85-104, 121-140

**FIGURE 59**

GGAAGGCAGCGGCAGCTCCACTCAGCCAGTACCCAGATACGCTGGGAACCTTCCCCAGCC**ATGGC**  
TTCCCTGGGGCAGATCCTCTTCTGGAGCATAATTAGCATCATCATTATTCTGGCTGGAGCAATTG  
CACTCATCATTGGCTTTGGTATTTTCAGGGAGACACTCCATCACAGTCACTACTGTGCCTCAGCT  
GGGAACATTGGGGAGGATGGAATCCTGAGCTGCACCTTTTGAACCTGACATCAAACCTTCTGATAT  
CGTGATACAATGGCTGAAGGAAGGTGTTTTAGGCTTGGTCCATGAGTTCAAAGAAGGCAAAGATG  
AGCTGTCTGGAGCAGGATGAAATGTTTCAGAGGCCGGACAGCAGTGTTTGCTGATCAAGTGATAGTT  
GGCAATGCCTCTTTGCGGCTGAAAAACGTGCAACTCACAGATGCTGGCACCTACAAATGTTATAT  
CATCACTTCTAAAGGCAAGGGGAATGCTAACCTTGAGTATAAACTGGAGCCTTCAGCATGCCGG  
AAGTGAATGTGGACTATAATGCCAGCTCAGAGACCTTGCGGTGTGAGGCTCCCCGATGGTTCCCC  
CAGCCCACAGTGGTCTGGGCATCCCAAGTTGACCAGGGAGCCAACCTTCTCGGAAGTCTCCAATAC  
CAGCTTTGAGCTGAACTCTGAGAATGTGACCATGAAGGTTGTGTCTGTGCTCTACAATGTTACGA  
TCAACAACACATACTCCTGTATGATTGAAAATGACATTGCCAAAGCAACAGGGGATATCAAAGTG  
ACAGAATCGGAGATCAAAAGGCGGAGTCACCTACAGCTGCTAAACTCAAAGGCTTCTCTGTGTGT  
CTCTTCTTTCTTTGCCATCAGCTGGGCACCTTCTGCCTCTCAGCCCTTACCTGATGCTAAAA**TAA**T  
GTGCCTTGGCCACAAAAAAGCATGCAAAGTCATTGTTACAACAGGGATCTACAGAACTATTTTAC  
CACCAGATATGACCTAGTTTTTATATTTCTGGGAGGAAATGAATTCATATCTAGAAGTCTGGAGTG  
AGCAAACAAGAGCAAGAAACAAAAAGAAGCCAAAAGCAGAAGGCTCCAATATGAACAAGATAAAT  
CTATCTTCAAAGACATATTAGAAGTTGGGAAAATAATTCATGTGAACTAGACAAGTGTGTTAAGA  
GTGATAAGTAAATGCACGTGGAGACAAGTGCATCCCCAGATCTCAGGGACCTCCCCCTGCCTGT  
CACCTGGGGAGTGAGAGGACAGGATAGTGCATGTTCTTTGTCTCTGAATTTTTTAGTTATATGTGC  
TGTAATGTTGCTCTGAGGAAGCCCCTGGAAAGTCTATCCCAACATATCCACATCTTATATTCCAC  
AAATTAAGCTGTAGTATGTACCCTAAGACGCTGCTAATTGACTGCCACTTCGCAACTCAGGGGCG  
GCTGCATTTTAGTAATGGGTCAAATGATTCACCTTTTTATGATGCTTCCAAAGGTGCCTTGGCTTC  
TCTTCCCAACTGACAAATGCCAAAGTTGAGAAAAATGATCATAATTTTAGCATAAACAGAGCAGT  
CGGGGACACCGATTTTATAAATAAACTGAGCACCTTCTTTTTTAAACAAAAAAAAAAAAAAAAAAAA  
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

## **FIGURE 60**

MASLGQILFWSIISIIIIILAGAIALIIGFGISGRHSITVTTVASAGNIGEDGILSCTFEPDIKLS  
DIVIQWLKEGVLGLVHEFKEGKDELSEQDEMFRGRTAVFADQVIVGNASRLKKNVQLTDAGTYKC  
YIITSKGKGNANLEYKTGAFSMPEVNVVDYNASSETLRCEAPRWFPOPTVVWASQVDQGANFSEVS  
NTSFELNSENVTMKVSVLYNVTINNTYSCMIENDIAKATGDIKVTESSEIKRRSHLQLLNSKASL  
CVSSFFAISWALLPLSPYLMLK

**Important features:**

**Signal peptide:**

amino acids 1-28

**Transmembrane domain:**

amino acids 258-281

**N-glycosylation sites.**

amino acids 112-116, 160-164, 190-194, 196-200, 205-209, 216-220,  
220-224

**N-myristoylation sites.**

amino acids 52-58, 126-132, 188-194

**FIGURE 61**

TGACGTCAGAATCACCATGGCCAGCTATCCTTACCGGCAGGGCTGCCCAGGAGCTGCAGGACAAG  
CACCAGGAGCCCCCTCCGGGTAGCTACTACCCTGGACCCCCCAATAGTGGAGGGCAGTATGGTAGT  
GGGCTACCCCCTGGTGGTGGTTATGGGGGTCTGCCCCTGGAGGGCCTTATGGACCACCAGCTGG  
TGGAGGGCCCTATGGACACCCCAATCCTGGGATGTTCCCCTCTGGAACCTCAGGAGGACCATATG  
GCGGTGCAGCTCCCGGGGGCCCCCTATGGTCAGCCACCTCCAAGTTCCTACGGTGCCCAGCAGCCT  
GGGCTTTATGGACAGGGTGGCGCCCCCTCCCAATGTGGATCCTGAGGCCTACTCCTGGTTCCAGTC  
GGTGGACTCAGATCACAGTGGCTATATCTCCATGAAGGAGCTAAAGCAGGCCCTGGTCAACTGCA  
ATTGGTCTTCATTCAATGATGAGACCTGCCTCATGATGATAAACATGTTTGACAAGACCAAGTCA  
GGCCGCATCGATGTCTACGGCTTCTCAGCCCTGTGGAAATTCATCCAGCAGTGGAAGAACCTCTT  
CCAGCAGTATGACCGGGACCGCTCGGGCTCCATTAGCTACACAGAGCTGCAGCAAGCTCTGTCCC  
AAATGGGCTACAACCTGAGCCCCCAGTTCACCCAGCTTCTGGTCTCCCGCTACTGCCCACGCTCT  
GCCAATCCTGCCATGCAGCTTGACCGCTTCATCCAGGTGTGCACCCAGCTGCAGGTGCTGACAGA  
GGCCTTCCGGGAGAAGGACACAGCTGTACAAGGCAACATCCGGCTCAGCTTCGAGGACTTCGTCA  
CCATGACAGCTTCTCGGATGCTATGAACCCAACCATCTGTGGAGAGTGGAGTGCACCAGGGACCTT  
TCCTGGCTTCTTAGAGTGAGAGAAGTATGTGGACATCTCTTCTTTTCTGTCCCTCTAGAAGAAC  
ATTCTCCCTTGCTTGATGCAACACTGTTCCAAAAGAGGGTGGAGAGTCCTGCATCATAGCCACCA  
AATAGTGAGGACCGGGGCTGAGGCCACACAGATAGGGGCCTGATGGAGGAGAGGATAGAAGTTGA  
ATGTCCTGATGGCCATGAGCAGTTGAGTGGCACAGCCTGGCACCAGGAGCAGGTCCTTGTAATGG  
AGTTAGTGTCCAGTCAGCTGAGCTCCACCCTGATGCCAGTGGTGAGTGTTTCATCGGCCTGTTACC  
GTTAGTACCTGTGTTCCCTCACCAGGCCATCCTGTCAAACGAGCCCATTTTCTCCAAAGTGGAAAT  
CTGACCAAGCATGAGAGAGATCTGTCTATGGGACCAGTGGCTTGATTCTGCCACACCCATAAAT  
CCTTGTGTGTAACTTCTAGCTGCCTGGGGCTGGCCCTGCTCAGACAAATCTGCTCCCTGGGCAT  
CTTTGGCCAGGCTTCTGCCCCCTGCAGCTGGGACCCCTCACTTGCCCTGCCATGCTCTGCTCGGCT  
TCAGTCTCCAGGAGACAGTGGTCACCTCTCCCTGCCAATACTTTTTTTAATTTGCATTTTTTTTC  
ATTTGGGGCCAAAAGTCCAGTGAAATTGTAAGCTTCAATAAAAGGATGAAACTCTGA

## **FIGURE 62**

MASYPYRQGCPGAAGQAPGAPPGSYYPGPPNSGGQYGSGLPPGGGYGGPAPGGPYGPPAGGGPYG  
HPNPGMFPSGTPGGPYGGAAPGGPYGQPPSSYGAQQPGLYGQGGAPPNVDPEAYSWFQSVDS DH  
SGYISMKELKQALVNCNWSSFNDETCLMMINMFDKTKSGRIDVYGFSALWKFIQQWKNLFQQYDR  
DRSGSISYTELQQALSQMGYNLSPQFTQLLVSRYCPRSANPAMQLDRFIQVCTQLQVLTEAFREK  
DTAVQGNIRLSFEDFVTMTASRML

**Important features of the protein:**

**Signal peptide:**

amino acids 1-19

**N-glycosylation site.**

amino acids 147-150

**Casein kinase II phosphorylation sites.**

amino acids 135-138, 150-153, 202-205, 271-274

**N-myristoylation sites.**

amino acids 9-14, 15-20, 19-24, 33-38, 34-39, 39-44, 43-48, 61-66,  
70-75, 78-83, 83-88, 87-92, 110-115



[illegible]

## **FIGURE 64**

MQGRVAGSCAPLGLLLVLHLPGLFARSIGVVEEKVSQNFGTNLPQLGQPSSTGPSNSEHPQPAL  
DPRSNDLARVPLKLSVPPSDGFPPAGGSAVQRWPPSWGLPAMDSWPPEDPWQMMAAAAEDRLGEA  
LPEELSYLSSAAALAPGSGPLPGESSPDATGLSPEASLLHQDSESRRLPRSNSLGAGGKILSQRP  
PWSLIHRVLPDHPWGTLNPSVSWGGGGPGTGWGTRPMPHPGEGIWGINNQPPGTSGWNINRYPGGS  
WGNINRYPGGSWGNINRYPGGSWGNIHLYPGINNPFPPGVLRPPGSSWNIPAGFPNPPSPRLQWG

**Important features of the protein:**

**Signal peptide:**

amino acids 1-26

**Casein kinase II phosphorylation sites.**

amino acids 56-59, 155-158

**N-myristoylation sites.**

amino acids 48-53, 220-225, 221-226, 224-229, 247-252, 258-263,  
259-264, 269-274, 270-275, 280-285, 281-286, 305-310

## **FIGURE 65**

AAGGAGAGGCCACCGGGACTTCAGTGTCTCCTCCATCCCAGGAGCGCAGTGGCCACTATGGGGTC  
TGGGCTGCCCCCTTGTCTCCTCTTGACCCTCCTTGGCAGCTCACATGGAACAGGGCCGGGTATGA  
CTTTGCAACTGAAGCTGAAGGAGTCTTTTCTGACAAATTCCTCCTATGAGTCCAGCTTCCTGGAA  
TTGCTTGAAAAGCTCTGCCTCCTCCTCCATCTCCCTTCAGGGACCAGCGTCACCCTCCACCATGC  
AAGATCTCAACACCATGTTGTCTGCAACACATGACAGCCATTGAAGCCTGTGTCCTTCTTGGCCC  
GGGCTTTTGGGCCGGGGATGCAGGAGGCAGGCCCCGACCCTGTCTTTCAGCAGGCCCCCACCCTC  
CTGAGTGGCAATAAATAAAATTCGGTATGCTG

## **FIGURE 66**

MGSGPLVLVLLLTLLGSSHGTGPGMTLQLKLKESFLTNSSESSFLELLEKLCLLLHLPSGTSVTL  
HHARSQHHVVCNT

**Important features:**

**Signal peptide:**

amino acids 1-19

**N-glycosylation site.**

amino acids 37-41

**N-myristoylation sites.**

amino acids 15-21, 19-25, 60-66

**FIGURE 67**

ACGGACCGAGGGTTCGAGGGAGGGACACGGACCAGGAACCTGAGCTAGGTCAAAGACGCCCCGGGC  
CAGGTGCCCCGTCGCAGGTGCCCCTGGCCGGAGATGCGGTAGGAGGGGCGAGCGCGAGAAGCCCC  
TTCCTCGGCGCTGCCAACCCGCCACCCAGCCCATGGCGAACCCCGGGCTGGGGCTGCTTCTGGCG  
CTGGGCCTGCCGTTCTTGCTGGCCCGCTGGGGCCGAGCCTGGGGGCAAATACAGACCACTTCTGC  
AAATGAGAATAGCACTGTTTTGCCTTCATCCACCAGCTCCAGCTCCGATGGCAACCTGCGTCCGG  
AAGCCATCACTGCTATCATCGTGGTCTTCTCCCTCTTGGCTGCCTTGCTCCTGGCTGTGGGGCTG  
GCACTGTTGGTGCGGAAGCTTCGGGAGAAGCGGCAGACGGAGGGCACCTACCGGCCCAGTAGCGA  
GGAGCAGTTCTCCCATGCAGCCGAGGCCCGGGCCCCCTCAGGACTCCAAGGAGACGGTGCAGGGCT  
GCCTGCCCATCTAGGTCCCCTCTCCTGCATCTGTCTCCCTTCATTGCTGTGTGACCTTGGGGAAA  
GGCAGTGCCCTCTCTGGGCAGTCAGATCCACCCAGTGCTTAATAGCAGGGAAGAAGGTACTTCAA  
AGACTCTGCCCCTGAGGTCAAGAGAGGATGGGGCTATTCACCTTTATATATTTATATAAAATTAG  
TAGTGAGATGTAAAAAAAAAAAAAAAAAAAA

## **FIGURE 68**

MANPGLGLLLALGLPFLRLRWGRAWGQIQTTSANENSTVLPSTSSSSDGNLRPEAITAIIVVFS  
LLAALLLAVGLALLVRKLREKRQTEGTYRPSSEEQFSHAAEARAPQDSKETVQGCLPI

**Important features:**

**Signal peptide:**

amino acids 1-19

**Transmembrane domain:**

amino acids 56-80

**N-glycosylation site.**

amino acids 36-40

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 86-90

**Tyrosine kinase phosphorylation site.**

amino acids 86-94

**N-myristoylation sites.**

amino acids 7-13, 26-32

### **FIGURE 69**

[illegible]

**FIGURE 70**

MGLFRGFVFLLVLCLLHQSNSTSFIKLNNNGFEDIVIVIDPSVPEDEKIIIEQIEDMVTTASTYLFE  
 ATEKRFFFKNVSILIPENWKENPQYKRPKHENHKHADVIVAPPTLPGRDEPYTKQFTECGEKGEY  
 IHFTPDLLLGGKKQNEYGPPGKLFVHEWAHLRWGVFDEYNEDQPFYRAKSKKIEATRCSAGISGRN  
 RVYKCQGGSCLSRACRIDSTTKLYGKDCQFFPDQVQTEKASIMFMQSIDSVVEFCNEKTHNQEAP  
 SLQNIKCNFRSTWEVISNSEDFKNTIPMVTPPPPPVFSLLKISQRIVCLVLDKSGSMGGKDRLNR  
 MNQAAKHFLQTVENGSWVGMVHFDSTATIVNKLIQIKSSDERNTLMAGLPTYPLGGTSICSGIK  
 YAFQVIGELHSQLDGSEVLLLLTDGEDNTASSCIDEVKQSGAIVHFIALGRAADEAVIEMSKITGG  
 SHFYVSDEAQNNGLIDAFGALTSGNTDLSQKSLQLESKGLTLNSNAWMNDTVIIDSTVGKDTFFL  
 ITWNSLPPSISLWDPSGTIMENFTVDATSKMAYLSIPGTAKVGTWAYNLQAKANPETLTITVTSR  
 AANSSVPPITVNAKMNKDVNSFPSPMIVYAEILQGYVPVLGANVTAFIESQNGHTEVLELLDNGA  
 GADSFKNMGVYSRYFTAYTENGRYSLKVRAHGGANTARLKLRPPLNRAAYIPGWVVNGEIEANPP  
 RPEIDEDTQTTLEDFSRTASGGAFVVSQVPSLPLPDQYPPSQITDLDATVHEDKIIILTWTA PGDN  
 FDVGKVQRYIIRISASILDLRDSFDDALQVNTTDLSPKEANSKESFAFKPENISEENATHIFIAI  
 KSIDKSNLTSKVSNIAQVTLFIPQANPDDIDPTPTPTPTPTPKSHNSGVNISTLVLSVIGSVVI  
 VNFILSTTI

**Signal peptide:**

amino acids 1-21

**Putative transmembrane domains:**

amino acids 284-300, 617-633

**Leucine zipper pattern.**

amino acids 469-491, 476-498

**N-glycosylation site.**

amino acids 20-24, 75-79, 340-344, 504-508, 542-546, 588-592,  
628-632, 811-815, 832-836, 837-841, 852-856, 896-900



**FIGURE 71**

CTCCTTAGGTTGGAAACCCTGGGAGTAGAGTACTGACAGCAAAGACCGGGAAAGACCATACGTCCTCCCGGGCAGGGGTGA  
CAACAGGTGTCATCTTTTTGATCTCGTGTGTGGCTGCCTTCCTATTTCAAGGAAAGACGCCAAGGTAATTTTGACCCA  
GAGGAGCAATGATGTAGCCACCTCCTAACCTTCCCTTCTTGAACCCCCAGTTATGCCAGGATTTACTAGAGAGTGTCA  
ACTCAACCAGCAAGCGGCTCCTTCGGCTTAACTTGTGGTTGGAGGAGAGAACCTTTGTGGGGCTGCGTTCTCTTAGCA  
GTGCTCAGAAGTGAAGTGCCTGAGGGTGGACCAGAAGAAAGGAAAGGTCCCCTCTTGCTGTTGGCTGCACATCAGGAA  
GGCTGTGATGGGAATGAAGGTGAAAACCTTGGAGATTTCACTTCAGTCATTGCTTCTGCCTGCAAGATCATCCTTTAAA  
AGTAGAGAAGCTGCTCTGTGTGGTGGTTAACTCCAAGAGGCAGAACTCGTTCTAGAAGGAAATGGATGCAAGCAGCTC  
CGGGGGCCCCAACGCATGCTTCTGTGGTCTAGCCCAGGGAAGCCCTTCCGTGGGGGGCCCCGGCTTTGAGGGATGCC  
ACCGGTTCTGGACGCATGGCTGATTCTGAATGATGATGGTTCCGCCGGGGGCTGCTTGCCTGGATTTCCTGGGTGGTG  
GTTTTGTGGTGTCTCTGTGTGCTATCTCTGTCTGTACATGTTGGCTGCACCCCAAGGTGACGAGGAGCAG  
CTGGCACTGCCCAGGGCCAACAGCCCCACGGGGAAGGAGGGGTACCAGGCCGTCCTTCAGGAGTGGGAGGAGCAGCAG  
CGCAACTACGTGAGCAGCCTGAAGCGGCAGATCGCACAGCTCAAGGAGGAGCTGCAGGAGAGGAGTGAAGCAGCTCAGG  
AATGGGCAGTACCAAGCCAGCGATGCTGCTGGCCTGGGTCTGGACAGGAGCCCCCAGAGAAAACCCAGGCCGACCTC  
CTGGCCTTCTGCACTCGCAGGTGGACAAGGCAGAGGTGAATGCTGGCGTCAAGCTGGCCACAGAGTATGCAGCAGTG  
CCTTTTCGATAGCTTTACTCTACAGAAGGTGTACCAGCTGGAGACTGGCCTTACCCGCCACCCCGAGGAGAAGCCTGTG  
AGGAAGGACAAGCGGGATGAGTTGGTGGAAAGCCATTGAATCAGCCTTGGAGACCCTGAACAATCTGCAGAGAACAGC  
CCCAATCACCGTCTTACACGGCCTCTGATTTTATAGAAGGGATCTACCGAACAGAAAGGGACAAAGGGACATTGTAT  
GAGCTCACCTTCAAAGGGGACCACAAACACGAATTCAAACGGCTCATCTTATTTTCGACCATTCAGCCCCATCATGAAA  
GTGAAAATGAAAAGCTCAACATGGCCAACACGCTTATCAATGTTATCGTGCCTCTAGCAAAAAGGGTGGACAAGTTC  
CGGCAGTTCATGCAGAATTTAGGGAGATGTGCATTGAGCAGGATGGGAGAGTCCATCTCACTGTTGTTTACTTTGGG  
AAAGAAGAAATAAATGAAGTCAAGGAATACTTGAAAACACTTCCAAAGCTGCCAACTTCAGGAACCTTTACCTTCATC  
CAGCTGAATGGAGAATTTTCTCGGGGAAGGGACTTGATGTTGGAGCCCGCTTCTGGAAGGGAAGCAACGTCTCTCTC  
TTTTTCTGTGATGTGGACATCTACTTCACATCTGAATTCCTCAATACGTGTAGGCTGAATACACAGCCAGGGAAGAAG  
GTATTTTATCCAGTTCCTTTTTCAGTCAGTACAATCCTGGCATAATATACGGCCACCATGATGCAGTCCCTCCCTTGAA  
CAGCAGCTGGTCATAAAGAAGGAACTGGATTTTGGAGAGACTTTGGATTTGGGATGACGTGTCAATATCGGTGAGC  
TTCATCAATATAGGTGGGTTTGATCTGGACATCAAAGGCTGGGGCGGAGAGGATGTGCACCTTTATCTCGCAATGATCTC  
CACAGCAACCTCATAGTGGTACGGACGCCTGTGCGAGGACTCTTCCACCTCTGGCATGAGAAGCGCTGCATGGACGAG  
CTGACCCCCGAGCAGTACAAGATGTGCATGCAGTCCAAGGCCATGAACGAGGCATCCACGGCCAGCTGGGCATGCTG  
GTGTTTCAGGCACGAGATAGAGGCTCACCTTCGCAACAGAAACAGAAGACAAGTAGCAAAAAACATGAAGTCCCTCCAG  
GAAGGATTGTGGGAGACACTTTTTCTTTCTTTTTCGCAATTACTGAAAGTGGCTGCAACAGAGAAAAGACTTCCATAAA  
GGACGACAAAAGAATTGGACTGATGGGTGAGAGATGAGAAAGCCTCCGATTTCTCTCTGTTGGGCTTTTTACAACAGA  
AATCAAAATCTCCGCTTTGCCTGCAAAAGTAACCCAGTTGCACCCTGTGAAGTGTCTGACAAAGGCAGAATGCTTGTG  
AGATTATAAGCCTAATGGTGTGGAGGTTTTGATGGTGTTTACAATACACTGAGACCTGTTGTTTTGTGTGCTCATTGA  
AATATTCATGATTTAAGAGCAGTTTTGTAAAAAATTCATTAGCATGAAAGGCAAGCATATTTCTCCTCATATGAATGA  
GCCTATCAGCAGGGCTCTAGTTTTCTAGGAATGCTAAATATCAGAAGGCAGGAGAGGAGATAGGCTTATTATGATACT  
AGTGAGTACATTAAGTAAAATAAAATGGACCAGAAAAGAAAAGAAACCATAAATATCGTGTCTATTTTTCCCAAGAT  
TAACCAAAAATAATCTGCTTATCTTTTTGGTTGTCTTTTAACTGTCTCCGTTTTTTTTCTTTTATTTAAAAATGCACT  
TTTTTTCCCTTGTGAGTTATAGTCTGCTTATTTAATTACCCTTTGCAAGCCTTACAAGAGAGCACAAGTTGGCCTAC  
ATTTTTATATTTTTTAAGAAGATACTTTGAGATGCATTATGAGAACTTTAGTTCAAAGCATCAAATTGATGCCATAT  
CCAAGGACATGCCAAATGCTGATTCTGTGAGGCACTGAATGTGAGGCATTGAGACATAGGGAAGGAATGGTTTTGTACT  
AATACAGACGTACAGATACTTTCTCTGAAGAGTATTTTCGAAGAGGAGCAACTGAACACTGGAGGAAAAGAAAATGAC  
ACTTTCTGCTTTACAGAAAAGGAACTCATTAGCATGGTGATATCGTGATGTACCTAAAAGTCAGAAACCACATTTT  
CTCCTCAGAAGTAGGGACCGCTTTCTTACCTGTTTTAAATAAACCAAAAGTATACCGTGTGAACCAAAACAATCTCTTTT  
AAAACAGGGTGCTCCTCCTGGCTTCTGGCTTCCATAAGAAGAAATGGAGAAAATATATATATATATATATATATTGT  
GAAAGATCAATCCATCTGCCAGAATCTAGTGGGATGGAAGTTTTTGTACATGTTATCCACCCAGGCCAGGTGGAAG  
TAAGTGAATTTATTTTTTAAATTAAGCAGTTCTACTCAATCACCAGATGCTTCTGAAAATTGCATTTTATTACCATTT  
CAAATATTTTTTAAAAATAAATACAGTTAATACATAGAGTGGTTTTCTTCATTATGTGAAAATTATTAGCCAGCACCAG  
ATGCATGAGCTAATTATCTCTTTGAGTCCTTGCTTCTGTTTGTCTCACAGTAACTCATTGTTTTAAAGCTTCAAGAAC  
ATTCAGCTGTTGGTGTGTTAAAAAATGCATTGTATTGATTTGTACTGGTAGTTTATGAAATTTAATTAAAAACACAGG  
CCATGAATGGAAGGTGGTATTGCACAGCTAATAAAATATGATTTGTGGATATGAA

## **FIGURE 72**

MMMVRRGLLAWISRVVLLVLLCCAISVLYMLACTPKGDEEQLALPRANSPTGKEGYQAVLQEW  
EQHRNYVSSSLKRQIAQLKEELQERSEQLRNGQYQASDAAGLGLDRSPPEKTQADLLAFLHSQVDK  
AEVNAGVKLATEYAAVPFDSFTLQKVYQLETGLTRHPEEKPVKDKRDELVEAIESALETLNNPA  
ENSPNHRPYTASDFIEGIYRTERDKGTLYELTFKGDHKHEFKRLILFRPFSPIMKVKNEKLNMAN  
TLINVIVPLAKRVDKFRQFMQNFREMCIEQDGRVHLTVVYFGKEEINEVKGILENTSKAANFRNF  
TFIQLNGEFSRGKGLDVGARFWKGSNVLLFFCDVDIYFTSEFLNTCRLNTQPGKKVFYPVLF  
SQYNPGIIYGHHDVAVPPLEQQQLVIKKETGFWRDFGFGMTCQYRSDFINIGGFDLDIKGWGGEDVHLYR  
KYLHSNLIIVRTPVRGLFHLWHEKRCMDELTPQYKMCMQSKAMNEASHGQLGMLVFRHEIEAHL  
RKQKQKTSSKKT

**Important features:**

**Signal peptide:**

amino acids 1-27

**N-glycosylation sites.**

amino acids 315-319, 324-328

**N-myristoylation sites.**

amino acids 96-102, 136-142, 212-218, 311-317, 339-345, 393-399

**Amidation site.**

amino acids 377-381

**FIGURE 73**

GAGACTGCAGAGGGAGATAAAGAGAGAGGGCAAAGAGGCAGCAAGAGATTTGTCCTGGGGATCCA  
GAAACCCATGATACCCTACTGAACACCGAATCCCCTGGAAGCCCACAGAGACAGAGACAGCAAGA  
GAAGCAGAGATAAATACACTCACGCCAGGAGCTCGCTCGCTCTCTCTCTCTCTCACTCCTC  
CCTCCCTCTCTCTCTGCCTGTCCTAGTCCTCTAGTCCTCAAATTCCCAGTCCCCTGCACCCCTTC  
CTGGGACACT**ATG**TTGTTCTCCGCCCTCCTGCTGGAGGTGATTTGGATCCTGGCTGCAGATGGGG  
GTCAACACTGGACGTATGAGGGCCCCACATGGTCAGGACCATTGGCCAGCCTCTTACCCTGAGTGT  
GGAAACAATGCCCAGTCGCCCATCGATATTCAGACAGACAGTGTGACATTTGACCCTGATTTGCC  
TGCTCTGCAGCCCCACGGATATGACCAGCCTGGCACCGAGCCTTTGGACCTGCACAACAATGGCC  
ACACAGTGCAACTCTCTCTGCCCTCTACCCTGTATCTGGGTGGACTTCCCCGAAAATATGTAGCT  
GCCAGCTCCACCTGCACTGGGGTCAGAAAGGATCCCCAGGGGGGTGAGAACACCAGATCAACAG  
TGAAGCCACATTTGCAGAGCTCCACATTGTACATTATGACTCTGATTCCTATGACAGCTTGAGTG  
AGGCTGCTGAGAGGCCTCAGGGCCTGGCTGTCTGGGCATCCTAATTGAGGTGGGTGAGACTAAG  
AATATAGCTTATGAACACATTCTGAGTCACTTGCATGAAGTCAGGCATAAAGATCAGAAGACCTC  
AGTGCCTCCCTTCAACCTAAGAGAGCTGCTCCCCAAACAGCTGGGGCAGTACTTCCGCTACAATG  
GCTCGCTCACAACCTCCCCCTTGCTACCAGAGTGTGCTCTGGACAGTTTTTTATAGAAGGTCCCAG  
ATTTCAATGGAACAGCTGGAAAAGCTTCAGGGGACATTGTTCTCCACAGAAGAGGAGCCCTCTAA  
GCTTCTGGTACAGAACTACCGAGCCCTTCAGCCTCTCAATCAGCGCATGGTCTTTGCTTCTTTCA  
TCCAAGCAGGATCCTCGTATACCACAGGTGAAATGCTGAGTCTAGGTGTAGGAATCTTGGTTGGC  
TGTCTCTGCCTTCTCCTGGCTGTTTATTTCAATTGCTAGAAAGATTCCGAAGAAGAGGCTGGAAAA  
CCGAAAGAGTGTGGTCTTCACCTCAGCACAAGCCACGACTGAGGCAT**TAA**ATTCCTTCTCAGATAC  
CATGGATGTGGATGACTTCCCTTCATGCCTATCAGGAAGCCTCTAAAATGGGGTGTAGGATCTGG  
CCAGAAACACTGTAGGAGTAGTAAGCAGATGTCCTCCTTCCCCTGGACATCTCTTAGAGAGGAAT  
GGACCCAGGCTGTCATTCCAGGAAGAACTGCAGAGCCTTCAGCCTCTCCAAACATGTAGGAGGAA  
ATGAGGAAATCGCTGTGTTGTTAATGCAGAGANCAAACCTCTGTTTAGTTGCAGGGGAAGTTTGGG  
ATATACCCCAAAGTCCTCTACCCCCTCACTTTTATGGCCCTTTCCCTAGATATACTGCGGGATCT  
CTCCTTAGGATAAAGAGTTGCTGTTGAAGTTGTATATTTTTGATCAATATATTTGGAAATTAAAG  
TTTCTGACTTT

## **FIGURE 74**

MLFSALLLEVIWILAADGGQHWTYEGPHGQDHWPA SYPECGNNAQSPIDIQTDSVTFDPDLPALQ  
PHGYDQPGTEPLDLHNNGHTVQLSLPSTLYLGGLPRKYVAAQLHLHWGQKGSPPGGSEHQINSEAT  
FAELHIVHYDSDSYDSLSEAAERPQGLAVLGILIEVGETKNIAYEHILSHLHEVRHKDQKTSVPP  
FNLRELLPKQLGQYFRYNGSLTTPPCYQSVLWTVFYRRSQISMEQLEKLQGTLFSTEEEPSKLLV  
QNYRALQPLNQRMVFASFIQAGSSYTTGEMLSLGVGILVGCLCLLLAVYFIARKIRKKRLENRKS  
VVFTSAQATTEA

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-15

#### **Transmembrane domain:**

amino acids 291-310

#### **N-glycosylation site.**

amino acids 213-216

#### **Eukaryotic-type carbonic anhydrases proteins**

amino acids 197-245, 104-140, 22-69

**FIGURE 75**

TGCCGCTGCCGCCGCTGCTGCTGTTGCTCCTGGCGGCGCCTTGGGGACGGGCAGTTCCCTGTGTC  
TCTGGTGGTTTGCCTAAACCTGCAAACATCACCTTCTTATCCATCAACATGAAGA**ATG**TCCTACA  
ATGGACTCCACCAGAGGGTCTTCAAGGAGTTAAAGTTACTTACACTGTGCAGTATTTTCATCACAA  
ATTGGCCCACCAGAGGTGGCACTGACTACAGATGAGAAGTCCATTTCTGTTGTCCTGACAGCTCC  
AGAGAAGTGGAAGAGAAATCCAGAAGACCTTCTGTTTCCATGCAACAAATATACTCCAATCTGA  
AGTATAACGTGTCTGTGTTGAATACTAAATCAAACAGAACGTGGTCCCAGTGTGTGACCAACCAC  
ACGCTGGTGCTCACCTGGCTGGAGCCGAACACTCTTTACTGCGTACACGTGGAGTCCTTCGTCCC  
AGGGCCCCCTCGCCGTGCTCAGCCTTCTGAGAAGCAGTGTGCCAGGACTTTGAAAGATCAATCAT  
CAGAGTTCAAGGCTAAAATCATCTTCTGGTATGTTTTGCCCATATCTATTACCGTGTTTCTTTTT  
TCTGTGATGGGCTATTCCATCTACCGATATATCCACGTTGGCAAAGAGAAACACCCAGCAAATTT  
GATTTTGATTTATGGAAATGAATTTGACAAAAGATTCTTTGTGCCTGCTGAAAAAATCGTGATTA  
ACTTTATCACCCCTCAATATCTCGGATGATTCTAAAATTTCTCATCAGGATATGAGTTTACTGGGA  
AAAAGCAGTGATGTATCCAGCCTTAATGATCCTCAGCCAGCGGGAACCTGAGGCCCCCTCAGGA  
GGAAGAGGAGGTGAAACATTTAGGGTATGCTTCGCATTTGATGGAAATTTTTTGTGACTCTGAAG  
AAAACACGGAAGGTACTTCTCTCACCCAGCAAGAGTCCCTCAGCAGAAACAATACCCCCGGATAAA  
ACAGTCATTGAATATGAATATGATGTCAGAACCACTGACATTTGTGCGGGGCCTGAAGAGCAGGA  
GCTCAGTTTGCAGGAGGAGGTGTCCACACAAGGAACATTATTGGAGTCGCAGGCAGCGTTGGCAG  
TCTTGGGCCCCGCAAACGTTACAGTACTCATAACCCCTCAGCTCCAAGACTTAGACCCCTGGCG  
CAGGAGCACACAGACTCGGAGGAGGGGCCGGAGGAAGAGCCATCGACGACCCTGGTCGACTGGGA  
TCCCCAAACTGGCAGGCTGTGTATTCTTCGCTGTCCAGCTTCGACCAGGATTAGAGGGCTGCG  
AGCCTTCTGAGGGGGATGGGCTCGGAGAGGAGGGTCTTCTATCTAGACTCTATGAGGAGCCGGCT  
CCAGACAGGCCACCAGGAGAAAAATGAAACCTATCTCATGCAATTCATGGAGGAATGGGGGTTATA  
TGTGCAGATGGAAAAC**TGA**TGCCAACACTTCCTTTTGCCTTTTGTTCCTGTGCAAACAAGTGAG  
TCACCCCTTTGATCCCAGCCATAAAGTACCTGGGATGAAAGAAGTTTTTTCCAGTTTGTGAGTGT  
CTGTGAGAATTACTTATTTCTTTTCTCTATTCTCATAGCACGTGTGTGATTGGTTCATGCATGTA  
GGTCTCTTAACAATGATGGTGGGCCTCTGGAGTCCAGGGGCTGGCCGGTTGTTCTATGCAGAGAA  
AGCAGTCAATAAATGTTTGCCAGACTGGGTGCAGAATTTATTAGGTGGGTGT

## **FIGURE 76**

MSYNGLHQRVFKELKLLTLCSSISQIGPPEVALTTDEKSISVVLTAPEKWKRNPEDLPVSMQQIY  
SNLKYNVSVLNTKSNRTWSQCVTNHTLVLTWLEPNTLYCVHVESFVPGPPRAQPSEKQCARTLK  
DQSSEFKAKIIFWYVLPISITVFLFSVMGYSIYRYIHVGKEKHPANLILYGNFEDKRFFVPAEK  
IVINFITLNISSDDSKISHQDMSLLGKSSDVSSLNDPQPSGNLRPPQEEEEVKHLGYASHLMEIFC  
DSEENTEGTSLTQQESLSRTIPDPKTVIEYEYDVRTTDCAGPEEQELSLQEEVSTQGTLLLESQA  
ALAVLGPQTLQYSYTPQLQDLPLAQEHTDSEEGPEEEPSTTLVDWDPQTGRLCIPSLSSFDQDS  
EGCEPSEG DGLGEEGLLSRLYEPPAPDRPPGENETYLMQFMEEWGLYVQMEN

**Important features:**

**Signal peptide:**

amino acids 1-28

**Transmembrane domain:**

amino acids 140-163

**N-glycosylation sites.**

amino acids 71-74, 80-83, 89-92, 204-207, 423-426

**FIGURE 77**

GAGGAGCGGGCCGAGGACTCCAGCGTGCCCAAGGTCTGGCATCCTGCACTTGCTGCCCTCTGACAC  
CTGGGAAGATGGCCGGCCCGTGGACCTTCACCCTTCTCTGTGGTTTGCTGGCAGCCACCTTGATC  
CAAGCCACCCTCAGTCCCCTGCACTTCTCATCCTCGGCCCAAAAGTCATCAAAGAAAAGCTGAC  
ACAGGAGCTGAAGGACCACAACGCCACCAGCATCCTGCAGCAGCTGCCGCTGCTCAGTGCCATGC  
GGGAAAAGCCAGCCGGAGGCATCCCTGTGCTGGGCAGCCTGGTGAACACCGTCCTGAAGCACATC  
ATCTGGCTGAAGGTCATCACAGCTAACATCCTCCAGCTGCAGGTGAAGCCCTCGGCCAATGACCA  
GGAGCTGCTAGTCAAGATCCCCCTGGACATGGTGGCTGGATTCAACACGCCCTGGTCAAGACCA  
TCGTGGAGTTCCACATGACGACTGAGGCCCAAGCCACCATCCGCATGGACACCAGTGCAAGTGGC  
CCCACCCGCCTGGTCCTCAGTGACTGTGCCACCAGCCATGGGAGCCTGCGCATCCAACTGCTGTA  
TAAGCTCTCCTTCTGGTGAACGCCTTAGCTAAGCAGGTCATGAACCTCCTAGTGCCATCCCTGC  
CCAATCTAGTGAAAAACCAGCTGTGTCCCGTGATCGAGGCTTCCTTCAATGGCATGTATGCAGAC  
CTCCTGCAGCTGGTGAAGGTGCCCATTTCCCTCAGCATTGACCGTCTGGAGTTTGACCTTCTGTA  
TCCTGCCATCAAGGGTGACACCATTAGCTCTACCTGGGGGCCAAGTTGTTGGACTCACAGGGAA  
AGGTGACCAAGTGGTTCAATAACTCTGCAGCTTCCCTGACAATGCCACCCTGGACAACATCCCG  
TTCAGCCTCATCGTGAGTCAGGACGTGGTGAAAGCTGCAGTGGCTGCTGTGCTCTCTCCAGAAGA  
ATTCATGGTCCTGTTGGACTCTGTGCTTCCTGAGAGTGCCCATCGGCTGAAGTCAAGCATCGGGC  
TGATCAATGAAAAGGCTGCAGATAAGCTGGGATCTACCCAGATCGTGAAGATCCTAACTCAGGAC  
ACTCCCGAGTTTTTTATAGACCAAGGCCATGCCAAGGTGGCCCAACTGATCGTGCTGGAAGTGTT  
TCCCTCCAGTGAAGCCCTCCGCCCTTTGTTACCCCTGGGCATCGAAGCCAGCTCGGAAGCTCAGT  
TTTACACCAAAGGTGACCAACTTATACTCAACTTGAATAACATCAGCTCTGATCGGATCCAGCTG  
ATGAACTCTGGGATTGGCTGGTTCCAACCTGATGTTCTGAAAAACATCATCACTGAGATCATCCA  
CTCCATCCTGCTGCCGAACCAGAATGGCAAATTAAGATCTGGGGTCCCAGTGTCATTGGTGAAGG  
CCTTGGGATTGAGGCAGCTGAGTCCTCACTGACCAAGGATGCCCTTGCTGCTTACTCCAGCCTCC  
TTGTGGAAACCCAGCTCTCCTGTCTCCAGTGAAGACTTGGATGGCAGCCATCAGGGAAGGCTGG  
GTCCAGCTGGGAGTATGGGTGTGAGCTCTATAGACCATCCCTCTCTGCAATCAATAAACACTTG  
CCTGTGAAAAA

## **FIGURE 78**

MAGPWTFTLLCGLLAATLIQATLSPTAVLILGPKVIKEKLTQELKDHNATSILQQLPLLSAMREK  
PAGGIPVLGSLVNTVLKHIIWLKVITANILQLQVKPSANDQELLVKIPLDMVAGFNTPLVKTIVE  
FHMTTEAQATIRMDTSASGPTRLVLSDCATSHGSLRIQLLYKLSFLVNALAKQVMNLLVPSLPNL  
VKNQLCPVIEASFNGMYADLLQLVKVPISLSIDRLEFDLLYPAlKGDTIQLYLGAKLLDSQGKVT  
KWFNNSAASLTMP TLDNIPFSLIVSQDVVKA AVAVLSPEEFMVLLDSVLPESA HRLKSSIGLIN  
EKAADKLGSTQIVKILTQDTPEFFIDQGHAKVAQLIVLEVFPSS EALRPLFTLGIEASSEAQFYT  
KGDQLILNLNNISSDRIQLMNSGIGWFQPDVLKNIITEIIHSILLPNQNGKLRSGVPVSLVKALG  
FEAAESSLT KDALVLTPASLWKPSSPVSQ

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-21

#### **N-glycosylation sites.**

amino acids 48-51, 264-267, 401-404

#### **Glycosaminoglycan attachment site.**

amino acids 412-415

#### **LBP / BPI / CETP family proteins.**

amino acids 407-457



**FIGURE 79**

GAGAGAAGTCAGCCTGGCAGAGAGACTCTGAAATGAGGGATTAGAGGTGTTCAAGGAGCAAGAGC  
TTCAGCCTGAAGACAAGGGAGCAGTCCCTGAAGACGCTTCTACTGAGAGGTCTGCC**ATG**GCCTCT  
CTTGGCCTCCAACTTGTGGGCTACATCCTAGGCCTTCTGGGGCTTTTGGGCACACTGGTTGCCAT  
GCTGCTCCCCAGCTGGAAAACAAGTTCTTATGTGGTGCCAGCATTGTGACAGCAGTTGGCTTCT  
CCAAGGGCCTCTGGATGGAATGTGCCACACACAGCACAGGCATCACCCAGTGTGACATCTATAGC  
ACCCTTCTGGGCCTGCCCCGCTGACATCCAGGCTGCCCAGGCCATGATGGTGACATCCAGTGCAAT  
CTCCTCCCTGGCCTGCATTATCTCTGTGGTGGGCATGAGATGCACAGTCTTCTGCCAGGAATCCC  
GAGCCAAAGACAGAGTGGCGGTAGCAGGTGGAGTCTTTTTCATCCTTGGAGGCCTCCTGGGATTC  
ATTCTGTTCCTGGAATCTTCATGGGATCCTACGGGACTTCTACTCACCCTGGTGCCTGACAG  
CATGAAATTTGAGATTGGAGAGGCTCTTTACTTGGGCATTATTTCTTCCCTGTTCTCCCTGATAG  
CTGGAATCATCCTCTGCTTTTCCTGCTCATCCCAGAGAAATCGCTCCAACCTACTACGATGCCTAC  
CAAGCCCAACCTCTTGCCACAAGGAGCTCTCCAAGGCCTGGTCAACCTCCCAAAGTCAAGAGTGA  
GTTCAATTCCTACAGCCTGACAGGGTATGTG**TGA**AGAACCAGGGGCCAGAGCTGGGGGGTGGCTG  
GGTCTGTGAAAAACAGTGGACAGCACCCCGAGGGCCACAGGTGAGGGACACTACCACTGGATCGT  
GTCAGAAGGTGCTGCTGAGGATAGACTGACTTTGGCCATTGGATTGAGCAAAGGCAGAAATGGGG  
GCTAGTGTAACAGCATGCAGGTTGAATTGCCAAGGATGCTCGCCATGCCAGCCTTTCTGTTTTCC  
TCACCTTGCTGCTCCCCTGCCCTAAGTCCCCAACCCCTCAACTTGAAACCCCATTCCTTAAGCCA  
GGACTCAGAGGATCCCTTTGCCCTCTGGTTTACCTGGGACTCCATCCCCAAACCCACTAATCACA  
TCCCACTGACTGACCCTCTGTGATCAAAGACCCTCTCTCTGGCTGAGGTTGGCTCTTAGCTCATT  
GCTGGGGATGGGAAGGAGAAGCAGTGGCTTTTGTGGGCATTGCTCTAACCTACTTCTCAAGCTTC  
CCTCCAAAGAACTGATTGGCCCTGGAACCTCCATCCCCTCTTGTTATGACTCCACAGTGTCCA  
GACTAATTTGTGCATGAACTGAAATAAAACCATCCTACGGTATCCAGGGAACAGAAAGCAGGATG  
CAGGATGGGAGGACAGGAAGGCAGCCTGGGACATTTAAAAAATA

## **FIGURE 80**

MASLGLQLVGYYLGLLGLLGLTLVAMLLPSWKTSSYVGASIVTAVGFSKGLWMECATHSTGITQCD  
IYSTLLGLPADIQAAQAMMVTSSAIISSLACIISVVGMRCTVFCQESRAKDRVAVAGGVFFILGGL  
LGFIPVAWNLHGILRDFYSPLVPDSMKFEIGEALYLGIISSLFSLIAGIILCFSCSSQRNRSNYY  
DAYQAQPLATRSSPRPGQPPKVKSEFNSYSLTGYV

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-24

#### **Transmembrane domains:**

amino acids 82-102, 117-140, 163-182

#### **N-glycosylation site.**

amino acids 190-193

#### **PMP-22 / EMP / MP20 family proteins.**

amino acids 46-59

**FIGURE 81**

CCCACGCGTCCGCGCCTCTCCCTTCTGCTGGACCTTCCTTCGTCTCTCCATCTCTCCCTCCTTTC  
CCCGCGTTCTCTTTCCACCTTTCTCTTCTTCCCACCTTAGACCTCCCTTCCTGCCCTCCTTTCTCCT  
GCCCACCGCTGCTTCCTGGCCCTTCTCCGACCCCGCTCTAGCAGCAGACCTCCTGGGGTCTGTGG  
GTTGATCTGTGGCCCCTGTGCCTCCGTGTCCTTTTCGTCTCCCTTCCTCCCGACTCCGCTCCCGG  
ACCAGCGGCCTGACCCTGGGGAAAGGATGGTTCCTGAGGTGAGGGTCTCTCCTCCTTGCTGGGA  
CTCGCGCTGCTCTGGTTCCCCCTGGACTCCACGCTCGAGCCCGCCAGACATGTTCTGCCTTTT  
CCATGGGAAGAGATACTCCCCGGCGAGAGCTGGCACCCCTACTTGAGGCCACAAGGCCTGATGT  
ACTGCCTGCGCTGTACCTGCTCAGAGGGCGCCCATGTGAGTTGTTACCGCCTCCACTGTCCGCCT  
GTCCACTGCCCCCAGCCTGTGACGGAGCCACAGCAATGCTGTCCCAAGTGTGTGGAACCTCACAC  
TCCCTCTGGACTCCGGGCCCCACCAAAGTCCTGCCAGCACAAACGGGACCATGTACCAACACGGAG  
AGATCTTCAGTGCCCATGAGCTGTTCCCTCCCGCCTGCCCAACCAGTGTGTCTCTGCAGCTGC  
ACAGAGGGCCAGATCTACTGCGGCCTCACAACTGCCCCGAACCAGGCTGCCCAGCACCCCTCCC  
ACTGCCAGACTCCTGCTGCCAAGCCTGCAAAGATGAGGCAAGTGAGCAATCGGATGAAGAGGACA  
GTGTGCAGTCGCTCCATGGGGTGAGACATCCTCAGGATCCATGTTCCAGTGATGCTGGGAGAAAG  
AGAGGCCCCGGGCACCCAGCCCCACTGGCCTCAGCGCCCCCTCTGAGCTTCATCCCTCGCCACTT  
CAGACCCAAGGGAGCAGGCAGCACAACTGTCAAGATCGTCCTGAAGGAGAAACATAAGAAAGCCT  
GTGTGCATGGCGGGAAGACGTACTCCACGGGGAGGTGTGGCACCCGGCCTTCCGTGCCTTCGGC  
CCCTTGCCCTGCATCCTATGCACCTGTGAGGATGGCCGCCAGGACTGCCAGCGTGTGACCTGTCC  
CACCGAGTACCCCTGCCGTACCCCGAGAAAGTGGCTGGGAAGTGTGCAAGATTTGCCCAGAGG  
ACAAAGCAGACCCTGGCCACAGTGAGATCAGTTCTACCAGGTGTCCAAGGCACCGGGCCGGGTC  
CTCGTCCACACATCGGTATCCCCAAGCCCAGACAACCTGCGTCGCTTTGCCCTGGAACACGAGGC  
CTCGGACTTGGTGGAGATCTACCTCTGGAAGCTGGTAAAAGATGAGGAACTGAGGCTCAGAGAG  
GTGAAGTACCTGGCCCAAGGCCACACAGCCAGAATCTTCCACTTGACTCAGATCAAGAAAGTCAG  
GAAGCAAGACTTCCAGAAAGAGGCACAGCACTTCCGACTGCTCGCTGGCCCCCACGAAGGTCCT  
GGAACGTCTTCTAGCCCAGACCCTGGAGCTGAAGGTCACGGCCAGTCCAGACAAAGTGACCAAG  
ACATAACAAAGACCTAACAGTTGCAGATATGAGCTGTATAATTGTTGTTATTATATATTAATAAA  
TAAGAAGTTGCATTACCCTCAAAAAAAAAAAAAAAAAAAAAA

## **FIGURE 82**

MVPEVRVLSSLLGLALLWFPLDSHARARPD MFCLFHGKRYSPGESWHPYLEPQGLMYCLRCTCSE  
GAHVSCYRLHCPPVHCPQPVTEPQQCCPKC VEPHTPSGLRAPPKSCQHNGTMYQHGEIFSAHELF  
PSRLPNQCVLCSCTEGQIYCGLTTCPEPGCPAPLPLPDSCCQACKDEASEQSDEEDSVQSLHGVR  
HPQDPCSSDAGRKRGP GTPAPTGLSAPLSFI PRHFRPKGAGSTTVKIVLKEKHKKACVHGGKTY S  
HGEVWHPAFRAFGPLPCILCTCEDGRQDCQRVTCPT EYPCRHPKVKAGKCKICPEDKADPGHSE  
ISSTRCPKAPGRVLVHTSVSPSPDNLRRFALEHEASDLVEIYLWKLVKDEETEAQRGEVPGPRPH  
SQNLPLDSDQESQEARLPERGTALPTARWP PRRSLERLPSPDPGAEGHGQSRQSDQDITKT

**Signal peptide:**

amino acids 1-25

**FIGURE 83**

GACAGCTGTGTCTCGATGGAGTAGACTCTCAGAACAGCGCAGTTTGCCCTCCGCTCACGCAGAGC  
CTCTCCGTGGCTTCCGCACCTTGAGCATTAGGCCAGTTCTCCTCTTCTCTCTAATCCATCCGTCA  
CCTCTCCTGTCATCCGTTTCCATGCCGTGAGGTCCATTACAGAACACATCCATGGCTCTCATGC  
TCAGTTTGGTTCTGAGTCTCCTCAAGCTGGGATCAGGGCAGTGGCAGGTGTTTGGGCCAGACAAG  
CCTGTCCAGGCCTTGGTGGGGGAGGACGCAGCATTCTCCTGTTTCTGTCTCCTAAGACCAATGC  
AGAGGCCATGGAAGTGCGGTTCTTCAGGGGCCAGTTCTCTAGCGTGGTCCACCTCTACAGGGACG  
GGAAGGACCAGCCATTTATGCAGATGCCACAGTATCAAGGCAGGACAAAACCTGGTGAAGGATTCT  
ATTGCGGAGGGGCGCATCTCTCTGAGGCTGGAAAACATTACTGTGTTGGATGCTGGCCTCTATGG  
GTGCAGGATTAGTTCCCAGTCTTACTACCAGAAGGCCATCTGGGAGCTACAGGTGTCAGCACTGG  
GCTCAGTTCCTCTCATTTCATCACGGGATATGTTGATAGAGACATCCAGCTACTCTGTCTAGTCC  
TCGGGCTGGTTCCCCCGGCCACAGCGAAGTGGAAGGTCCACAAGGACAGGATTTGTCCACAGA  
CTCCAGGACAAACAGAGACATGCATGGCCTGTTTGATGTGGAGATCTCTCTGACCGTCCAAGAGA  
ACGCCGGGAGCATATCCTGTTCCATGCGGCATGCTCATCTGAGCCGAGAGGTGGAATCCAGGGTA  
CAGATAGGAGATACTTTTTTCGAGCCTATATCGTGGCACCTGGCTACCAAAGTACTGGGAATACT  
CTGCTGTGGCCTATTTTTTGGCATTGTTGGACTGAAGATTTTCTTCTCCAAATTCAGTGGAAAA  
TCCAGGCGGAACTGGACTGGAGAAGAAAGCACGGACAGGCAGAATTGAGAGACGCCCGGAAACAC  
GCAGTGGAGGTGACTCTGGATCCAGAGACGGCTCACCCGAAGCTCTGCGTTTCTGATCTGAAAAC  
TGTAACCCATAGAAAAGCTCCCCAGGAGGTGCCTCACTCTGAGAAGAGATTTACAAGGAAGAGTG  
TGGTGGCTTCTCAGAGTTTCCAAGCAGGGAAACATTACTGGGAGGTGGACGGAGGACACAATAAA  
AGGTGGCGCGTGGGAGTGTGCCGGGATGATGTGGACAGGAGGAAGGAGTACGTGACTTTGTCTCC  
CGATCATGGGTACTGGGTCCTCAGACTGAATGGAGAACATTTGTATTTACATTAAATCCCCGTT  
TTATCAGCGTCTTCCCCAGGACCCACCTACAAAAATAGGGGTCTTCTGGACTATGAGTGTGGG  
ACCATCTCCTTCTTCAACATAAATGACCAGTCCCTTATTTATACCCTGACATGTCGGTTTGAAGG  
CTTATTGAGGCCCTACATTGAGTATCCGTCTATAATGAGCAAAATGGAATCCCATAGTCATCT  
GCCCAGTCACCCAGGAATCAGAGAAAGAGGCCTCTTGGCAAAGGGCCTCTGCAATCCCAGAGACA  
AGCAACAGTGAGTCCTCCTCACAGGCAACCACGCCCTTCTCCTCCCCAGGGGTGAAATGTAGGATGA  
ATCACATCCACATTCTTCTTTAGGGATATTAAGGTCTCTCTCCCAGATCCAAAGTCCCGCAGCA  
GCCGGCCAAGGTGGCTTCCAGATGAAGGGGGACTGGCCTGTCCACATGGGAGTCAGGTGTCATGG  
CTGCCCTGAGCTGGGAGGGAAGAAGGCTGACATTACATTTAGTTTGCTCTCACTCCATCTGGCTA  
AGTGATCTTGAAATACCACCTCTCAGGTGAAGAACCGTCAGGAATTCCCATCTCACAGGCTGTGG  
TGTAAGATTAAGTAGACAAGGAATGTGAATAATGCTTAGATCTTATTGATGACAGAGTGTATCCTA  
ATGGTTTGTTTATTATATTACACTTTCAGTAAAAAA

## **FIGURE 84**

MALMLSLVLSLLKLGGQWQVFGPDKPVQALVGEDAAFSCFLSPKTNAEAMEVRFFRGQFSSVH  
LYRDGKDQPFMQMPQYQGRTKLVKDSIAEGRISLRLENITVLDAGLYGCRISSQSYQKAIWELQ  
VSALGSVPLISITGYVDRDIQLLCQSSGWFPRTAKWKGPQGQDLSTDSRTNRDMHGLFDVEISL  
TVQENAGSISCSMRHAHLSREVESRVQIGDTFFEPISWHLATKVLGILCCGLFFGIVGLKIFFSK  
FQWKIQAELDWRRKHGQAEIRDARKHAVEVTLDPETAHPKLCVSDLKTVTHRKAPQEVPHSEKRF  
TRKSVVASQS FQAGKHYWEVDGGHNKRWRVGVCRDDVDRKEYVTLSPDHGYWVLRNLNGEHLFT  
LNPRFISVFPRTPPTKIGVFLDYECGTISFFNINDQSLIYTLTCRFEGLLRPYIEYPSYNEQNGT  
PIVICPVTQESEKEASWQRASAI PETSNSSESSSQATTPFLPRGEM

**Signal peptide:**

amino acids 1-17

**Transmembrane domain:**

amino acids 239-255

**FIGURE 85**

AACAGACGTTCCCTCGCGGCCCTGGCACCTCTAACCCCAGACATGCTGCTGCTGCTGCTGCCCCCT  
GCTCTGGGGGAGGGAGAGGGCGGAAGGACAGACAAGTAACTGCTGACGATGCAGAGTTCCGTGA  
CGGTGCAGGAAGGCCTGTGTGTCCATGTGCCCTGCTCCTTCTCCTACCCCTCGCATGGCTGGATT  
TACCCTGGCCCAGTAGTTCATGGCTACTGGTTCCGGGAAGGGGCCAATACAGACCAGGATGCTCC  
AGTGGCCACAAACAACCCAGCTCGGGCAGTGTGGGAGGAGACTCGGGACCGATTCCACCTCCTTG  
GGGACCCACATACCAAGAATTGCACCCTGAGCATCAGAGATGCCAGAAGAAGTGATGCGGGGAGA  
TACTTCTTTCGTATGGAGAAAGGAAGTATAAAATGGAATTATAAACATCACCGGCTCTCTGTGAA  
TGTGACAGCCTTGACCCACAGGCCCAACATCCTCATCCCAGGCACCCTGGAGTCCGGCTGCCCCC  
AGAATCTGACCTGCTCTGTGCCCTGGGCCTGTGAGCAGGGGACACCCCCTATGATCTCCTGGATA  
GGGACCTCCGTGTCCCCCTGGACCCCTCCACCACCCGCTCCTCGGTGCTCACCTCATCCCACA  
GCCCCAGGACCATGGCACCAAGCCTCACCTGTCAGGTGACCTTCCCTGGGGCCAGCGTGACCACGA  
ACAAGACCGTCCATCTCAACGTGTCTACCCGCCTCAGAACTTGACCATGACTGTCTTCCAAGGA  
GACGGCACAGTATCCACAGTCTTGGGAAATGGCTCATCTCTGTCACTCCCAGAGGGCCAGTCTCT  
GCGCCTGGTCTGTGCAGTTGATGCAGTTGACAGCAATCCCCCTGCCAGGCTGAGCCTGAGCTGGA  
GAGGCCTGACCCTGTGCCCTCACAGCCCTCAAACCCGGGGGTGCTGGAGCTGCCTTGGGTGCAC  
CTGAGGGATGCAGCTGAATTCACCTGCAGAGCTCAGAACCTCTCGGCTCTCAGCAGGTCTACCT  
GAACGTCTCCCTGCAGAGCAAAGCCACATCAGGAGTGACTCAGGGGGTGGTGGGGGAGCTGGAG  
CCACAGCCCTGGTCTTCCTGTCCTTCTGCGTCATCTTCGTTGTAGTGAGGTCCTGCAGGAAGAAA  
TCGGCAAGGCCAGCAGCGGGCGTGGGAGATACGGGCATAGAGGATGCAAACGCTGTCAGGGGTTC  
AGCCTCTCAGGGGCCCCTGACTGAACCTTGGGCAGAAGACAGTCCCCCAGACCAGCCTCCCCAG  
CTTCTGCCCCTCCTCAGTGGGGGAAGGAGAGCTCCAGTATGCATCCCTCAGCTTCCAGATGGTG  
AAGCCTTGGGACTCGCGGGGACAGGAGGCCACTGACACCGAGTACTCGGAGATCAAGATCCACAG  
ATGAGAAACTGCAGAGACTCACCTGATTGAGGGATCACAGCCCCTCCAGGCAAGGGAGAAGTCA  
GAGGCTGATTCTTGTAGAATTAACAGCCCTCAACGTGATGAGCTATGATAACACTATGAATTATG  
TGCAGAGTGAAAAGCACACAGGCTTTAGAGTCAAAGTATCTCAAACCTGAATCCACACTGTGCCC  
TCCCTTTTATTTTTTTAACTAAAAGACAGACAAATTCCTA

## **FIGURE 86**

MLLLLLPLLWGRERAEGQTSKLLTMQSSVTVQEGLCVHVPCSFSSYP SHGWIYPGPVVHGYWFREG  
ANTDQDAPVATNNPARAVWEETRDRFHLLGDPHTKNCTLSIRDARRSDAGRYFFRMEKGSIKWNY  
KHHRLSVNVTALTHRPNILIPGTLESGCPQNLTCVWPWACEQGTPPMISWIGTSVSPLDPSTTRS  
SVLTLLIPQPQDHGTSLTCQVTFPGASVTTNKTVHLNVSYPPQNLTMTVFQGDGTVSTVLGNGSSL  
SLPEGQSLRLVCAVDAVDSNPPARLSLSWRGLTLCPSQPSNPGVLELPWVHLRDAAEFTCRAQNP  
LGSQQVYLVNLSLQSKATSGVTQGVVGGAGATALVFLSFCVIFVVVRSCRKKSARPAAGVGD TGIE  
DANAVRG SASQG PLTEPWAEDSPPDQPPPASARSSVGE GELQYASLSFQMVKPWDSRGQEATDTE  
YSEIKIHR

**Signal peptide:**

amino acids 1-15

**Transmembrane domain:**

amino acids 351-370



**FIGURE 87**

AGAAAGCTGCACTCTGTTGAGCTCCAGGGCGCAGTGGAGGGAGGGAGTGAAGGAGCTCTCTGTAC  
CCAAGGAAAGTGCAGCTGAGACTCAGACAAGATTACA**ATG**AACCAACTCAGCTTCCTGCTGTTTC  
TCATAGCGACCACCAGAGGATGGAGTACAGATGAGGCTAATACTTACTTCAAGGAATGGACCTGT  
TCTTCGTCTCCATCTCTGCCCAGAAGCTGCAAGGAAATCAAAGACGAATGTCCTAGTGCATTTGA  
TGGCCTGTATTTTCTCCGCACTGAGAATGGTGTATCTACCAGACCTTCTGTGACATGACCTCTG  
GGGTGGCGGCTGGACCCTGGTGGCCAGCGTGCATGAGAATGACATGCGTGGGAAGTGCACGGTG  
GGCGATCGCTGGTCCAGTCAGCAGGGCAGCAAAGCAGACTACCCAGAGGGGGACGGCAACTGGGC  
CAACTACAACACCTTTGGATCTGCAGAGGCGGCCACGAGCGATGACTACAAGAACCCTGGCTACT  
ACGACATCCAGGCCAAGGACCTGGGCATCTGGCACGTGCCCAATAAGTCCCCCATGCAGCACTGG  
AGAAACAGCTCCCTGCTGAGGTACCGCACGGACACTGGCTTCCTCCAGACACTGGGACATAATCT  
GTTTGGCATCTACCAGAAATATCCAGTGAAATATGGAGAAGGAAAGTGTTGGACTGACAACGGCC  
CGGTGATCCCTGTGGTCTATGATTTTGGCGACGCCCAGAAAACAGCATCTTATTACTCACCCCTAT  
GGCCAGCGGGAATTCACTGCGGGATTTGTTCA GTTCAGGGTATTTAATAACGAGAGAGCAGCCAA  
CGCCTTGTGTGCTGGAATGAGGGTCACCGGATGTAACACTGAGCATCACTGCATTGGTGGAGGAG  
GATACTTTCCAGAGGCCAGTCCCCAGCAGTGTGGAGATTTTCTGGTTTTGATTGGAGTGGATAT  
GGA ACTCATGTTGGTTACAGCAGCAGCCGTGAGATAACTGAGGCAGCTGTGCTTCTATTCTATCG  
T**TGA**GAGTTTTGTGGGAGGGAACCCAGACCTCTCCTCCCAACCATGAGATCCCAAGGATGGAGAA  
CAACTTACCCAGTAGCTAGAATGTTAATGGCAGAAGAGAAAACAATAAATCATATTGACTCAAGA  
AAAAAA

## **FIGURE 88**

MNQLSFLLFLIATTRGWSTDEANTYFKEWTCSSSPSLPRSCKEIKDECPSAFDGLYFLRTENGVI  
YQTFCDMTSGGGGWTLVASVHENDMRGKCTVGDRWSSQQGSKADYPEGDGNWANYNTFGSAEAAT  
SDDYKNPGYYDIQAKDLGIWHVPNKSPMQHWRNSSLLRYRTDTGFLQTLGHNLFGIYQKYPVKYG  
EGKCWTDNGPVI PVVYDFGDAQKTASYSPYGQREFTAGFVQFRVFNNERAANALCAGMRVTGCN  
TEHHCIGGGGYFPEASPPQCGDFSGFDWSGYGTHVGYSSSREITEAAVLLFYR

**Important features:**

**Signal peptide:**

amino acids 1-16

**N-glycosylation site.**

amino acids 163-167

**Glycosaminoglycan attachment sites.**

amino acids 74-78, 289-293

**N-myristoylation sites.**

amino acids 76-82, 115-121, 124-130, 253-259, 292-298

**FIGURE 89**

CTAGATTTGTCGGCTTGCGGGGAGACTTCAGGAGTCGCTGTCTCTGAACTTCCAGCCTCAGAGAC  
CGCCGCCCTTGTCCCCGAGGGGCC**ATG**GGCCGGGTCTCAGGGCTTGTGCCCTCTCGCTTCCTGACG  
CTCCTGGCGCATCTGGTGGTCGTCATCACCTTATTCTGGTCCCGGGACAGCAACATACAGGCCTG  
CCTGCCTCTCACGTTACCCCCGAGGAGTATGACAAGCAGGACATTCAGCTGGTGGCCGCGCTCT  
CTGTCACCCTGGGCCTCTTTGCAGTGGAGCTGGCCGGTTTCCTCTCAGGAGTCTCCATGTTCAAC  
AGCACCCAGAGCCTCATCTCCATTGGGGCTCACTGTAGTGCATCCGTGGCCCTGTCCTTCTTCAT  
ATTCGAGCGTTGGGAGTGCACCTACGTATTGGTACATTTTTGTCTTCTGCAGTGCCCTTCCAGCTG  
TCACTGAAATGGCTTTATTCGTCACCGTCTTTGGGCTGAAAAGAAACCCTTC**TGA**TTACCTTCA  
TGACGGGAACCTAAGGACGAAGCCTACAGGGGCAAGGGCCGCTTCGTATTCCTGGAAGAAGGAAG  
GCATAGGCTTCGGTTTTCCCCTCGGAAACTGCTTCTGCTGGAGGATATGTGTTGGAATAATTACG  
TCTTGAGTCTGGGATTATCCGCATTGTATTTAGTGCTTTGTAATAAAATATGTTTTGTAGTAACA  
TTAAGACTTATATACAGTTTTAGGGGACAATTAAAAAAAAAAAA

## **FIGURE 90**

MGRVSGLVPSRFLTLLAHLVVVITLFWSRDSNIQACLPLTFTPEEYDKQDIQLVAALSVTLGLFA  
VELAGFLSGVSMFNSTQSLISIGAHCSASVALSFFIFERWECTTYWYIFVFCSALPAVTEMALFV  
TVFGLKKKPF

**Transmembrane domain:**

amino acids 12-28 (type II), 51-66, 107-124

**FIGURE 91**

CTGGGACCCCGAAAAGAGAAGGGGAGAGCGAGGGGACGAGAGCGGAGGAGGAAGATGCAACTGAC  
TCGCTGCTGCTTCGTGTTCTTGGTGCAGGGTAGCCTCTATCTGGTCATCTGTGGCCAGGATGATG  
GTCCTCCCGGCTCAGAGGACCCTGAGCGTGATGACCACGAGGGCCAGCCCCGGCCCCGGGTGCCT  
CGGAAGCGGGGCCACATCTCACCTAAGTCCCGCCCCATGGCCAATTCCACTCTCCTAGGGCTGCT  
GGCCCCGCCTGGGGAGGCTTGGGGCATTCTTGGGCAGCCCCCAACCGCCCCGAACCACAGCCCCC  
CACCTCAGCCAAGGTGAAGAAAATCTTTGGCTGGGGCGACTTCTACTCCAACATCAAGACGGTG  
GCCCTGAACCTGCTCGTCACAGGGAAGATTGTGGACCATGGCAATGGGACCTTCAGCGTCCACTT  
CCAACACAATGCCACAGGCCAGGGAAACATCTCCATCAGCCTCGTGCCCCCAGTAAAGCTGTAG  
AGTTCCACCAGGAACAGCAGATCTTCATCGAAGCCAAGGCCTCCAAAATCTTCAACTGCCGGATG  
GAGTGGGAGAAGGTAGAACGGGGCCGCCGACCTCGCTTTGCACCCACGACCCAGCCAAGATCTG  
CTCCCGAGACCACGCTCAGAGCTCAGCCACCTGGAGCTGCTCCCAGCCCTTCAAAGTCGTCTGTG  
TCTACATCGCCTTCTACAGCACGGACTATCGGCTGGTCCAGAAGGTGTGCCCAGATTACAACCTAC  
CATAGTGATACCCCCTACTACCCATCTGGGTGACCCGGGGCAGGCCACAGAGGCCAGGCCAGGGC  
TGGAAGGACAGGCCTGCCCATGCAGGAGACCATCTGGACACCGGGCAGGGAAGGGGTGGGCCTC  
AGGCAGGGAGGGGGGTGGAGACGAGGAGATGCCAAGTGGGGCCAGGGCCAAGTCTCAAGTGGCAG  
AGAAAGGGTCCCAAGTGCTGGTCCCAACCTGAAGCTGTGGAGTGACTAGATCACAGGAGCACTGG  
AGGAGGAGTGGGCTCTCTGTGCAGCCTCAGAGGGCTTGGCCACGGAGCCACAGAGAGATGCTGGG  
TCCCGAGGCCTGTGGGCAGGCCGATCAGTGTGGCCCCAGATCAAGTCATGGGAGGAAGCTAAGC  
CCTTGGTTCTTGCCATCCTGAGGAAAGATAGCAACAGGGAGGGGGAGATTTTCATCAGTGTGGACA  
GCCTGTCAACTTAGGATGGATGGCTGAGAGGGCTTCTTAGGAGCCAGTCAGCAGGGTGGGGTGGG  
GCCAGAGGAGCTCTCCAGCCCTGCCTAGTGGGCGCCCTGAGCCCCTTGTCGTGTGCTGAGCATGG  
CATGAGGCTGAAGTGGCAACCCTGGGGTCTTTGATGTCTTGACAGATTGACCATCTGTCTCCAGC  
CAGGCCACCCCTTTCCAAAATTCCCTCTTCTGCCAGTACTCCCCCTGTACCACCCATTGCTGATG  
GCACACCCATCCTTAAGCTAAGACAGGACGATTGTGGTCCTCCACACTAAGGCCACAGCCCATC  
CGCGTGCTGTGTGTCCCTCTTCCACCCCCAACCCCTGCTGGCTCCTCTGGGAGCATCCATGTCCCG  
GAGAGGGGTCCCTCAACAGTCAGCCTCACCTGTCAGACCGGGGTCTCCCGGATCTGGATGGCGC  
CGCCCTCTCAGCAGCGGGCACGGGTGGGGCGGGGCCGGGCCGAGAGCATGTGCTGGATCTGTTC  
TGTGTGTCTGTCTGTGGGTGGGGGAGGGGAGGGAAGTCTTGTGAAACCGCTGATTGCTGACTTT  
TGTGTGAAGAATCGTGTTCTTGGAGCAGGAAATAAAGCTTGCCCCGGGGCA

## **FIGURE 92**

MQLTRCCFVFLVQGSLYLVICGQDDGPPGSEDPERDDHEGQPRPRVPRKRGHISPKSRPMANSTL  
LGLLAPPGEAWGILGQPPNRPNHSPPPSAKVKKIFGWGDFYSNIKTVALNLLVTGKIVDHGNGTF  
SVHFQHNATGQGNISISLVPPSKAVEFHQEQQIFIEAKASKIFNCRMWEKVERGRRTSLCTHDP  
AKICSRDHAQSSATWSCSQPFKVVCVYIAFYSTDYRLVQKVCPDYNHSDTPYYPSG

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-14

#### **N-glycosylation sites.**

amino acids 62-65, 127-130, 137-140, 143-146

#### **2-oxo acid dehydrogenases acyltransferase**

amino acids 61-71

**FIGURE 93**

CGGTGGCCATGACTGCGGCCGTGTTCTTCGGCTGCGCCTTCATTGCCTTCGGGCCTGCGCTCGCC  
CTTTATGTCTTCACCATCGCCATCGAGCCGTTGCGTATCATCTTCCTCATCGCCGGAGCTTTCTT  
CTGGTTGGTGTCTCTACTGATTTTCGTCCCTTGTTTGGTTCATGGCAAGAGTCATTATTGACAACA  
AAGATGGACCAACACAGAAATATCTGCTGATCTTTGGAGCGTTTGTCTCTGTCTATATCCAAGAA  
ATGTTCCGATTTGCATATTATAAACTCTTAAAAAAGCCAGTGAAGGTTTGAAGAGTATAAACCC  
AGGTGAGACAGCACCCCTCTATGCGACTGCTGGCCTATGTTTCTGGCTTGGGCTTTGGAATCATGA  
GTGGAGTATTTTCCTTTGTGAATACCCTATCTGACTCCTTGGGGCCAGGCACAGTGGGCATTCAT  
GGAGATTCTCCTCAATTCTTCCTTTATTCAGCTTTCATGACGCTGGTCATTATCTTGCTGCATGT  
ATTCTGGGGCATTGTATTTTTTTGATGGCTGTGAGAAGAAAAAGTGGGGCATCCTCCTTATCGTTC  
TCCTGACCCACCTGCTGGTGTGAGCCAGACCTTCATAAGTTCTTATTATGGAATAAACCTGGCG  
TCAGCATTTATAATCCTGGTGCTCATGGGCACCTGGGCATTCTTAGCTGCGGGAGGCAGCTGCCG  
AAGCCTGAAACTCTGCCTGCTCTGCCAAGACAAGAACTTTCTTCTTTACAACCAGCGCTCCAGATT  
AACCTCAGGGAACCAGCACTTCCCAAACCGCAGACTACATCTTTAGAGGAAGCACAACTGTGCCT  
TTTTCTGAAAATCCCTTTTTCTGGTGGAATTGAGAAAGAAATAAACTATGCAGATA

## **FIGURE 94**

MTAAVFFGCAFIAGFPALALYVFTIAIEPLRIIFLIAGAFFWLVSLLISSLVWFMARVIIDNKDG  
PTQKYLLIFGAFVSVYIQEMFRFAYYKLLKKASEGLKSINPGETAPSMRLLAYVSGLGFGIMSGV  
FSFVNTLSDSLGPSTVGIHGDSPQFFLYSAFMTLVIILLHVFWGIVFFDGCEKKKWGILLIVLLT  
HLLVSAQTFISSYYGINLASAFIILVLMGTWAFLAAGGSCRSLKLCLLCQDKNFLLYNQSR

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-19

#### **Transmembrane domains:**

amino acids 32-51, 119-138, 152-169, 216-235

#### **Glycosaminoglycan attachment site.**

amino acids 120-123

#### **Sodium:neurotransmitter symporter family protein**

amino acids 31-65



**FIGURE 95**

AATTTTTCACCAGAGTAAACTTGAGAAACCAACTGGACCTTGAGTATTGTACATTTTGCCTCGTG  
GACCCAAAGGTAGCAATCTGAAACATGAGGAGTACGATTCTACTGTTTTGTCTTCTAGGATCAAC  
TCGGTCATTACCACAGCTCAAACCTGCTTTGGGACTCCCTCCCACAAAACCTGGCTCCGGATCAGG  
GAACACTACCAAACCAACAGCAGTCAAATCAGGTCTTTCCTTCTTTAAGTCTGATACCATTAACA  
CAGATGCTCACACTGGGGCCAGATCTGCATCTGTTAAATCCTGCTGCAGGAATGACACCTGGTAC  
CCAGACCCACCCATTGACCCTGGGAGGGTTGAATGTACAACAGCAACTGCACCCACATGTGTTAC  
CAATTTTTGTACACAACCTGGAGCCCAGGGCACTATCCTAAGCTCAGAGGAATTGCCACAAATC  
TTCACGAGCCTCATCATCCATTCCCTTGTTCCCGGGAGGCATCCTGCCCACCAGTCAGGCAGGGGC  
TAATCCAGATGTCCAGGATGGAAGCCTTCCAGCAGGAGGAGCAGGTGTAAATCCTGCCACCCAGG  
GAACCCAGCAGGCCGCCTCCCAACTCCCAGTGGCACAGATGACGACTTTGCAGTGACCACCCCT  
GCAGGCATCCAAAGGAGCACACATGCCATCGAGGAAGCCACCACAGAATCAGCAAATGGAATTCA  
GTAAGCTGTTTCAAATTTTTTCAACTAAGCTGCCTCGAATTTGGTGATACATGTGAATCTTTATC  
ATTGATTATATTATGGAATAGATTGAGACACATTGGATAGTCTTAGAAGAAATTAATTCTTAATT  
TACCTGAAAATATTCTTGAAATTCAGAAAATATGTTCTATGTAGAGAATCCCAACTTTTAAAAA  
CAATAATTCAATGGATAAATCTGTCTTTGAAATATAACATTATGCTGCCTGGATGATATGCATAT  
TAAAACATATTTGGAAAACCTGGAAA  
AAA

## **FIGURE 96**

MRSTILLFCLLGSTRSLPQLKPALGLPPTKLAPDQGTLPNQQQSNQVFPSLSLIPLTQM  
LTLGPDHLHLLNPAAGMTPGTQTHPLTLGGLNVQQQLHHPHVLPIFVTQLGAQGTLISSEE  
LPQIFTSLIIHSLEFPGGILPTSQAGANPDVQDGSLPAGGAGVNPATQGTPAGRLPTPSG  
TDDDFAVTTPAGIQRSTHAIEEATTESANGIQ

**Signal peptide:**

amino acids 1-16

**FIGURE 97**

GCTCAAGTGCCCTGCCTTGCCCCACCCAGCCAGCCTGGCCAGAGCCCCCTGGAGAAGGAGCTCT  
 CTTCTTGCTTGGCAGCTGGACCAAGGGAGCCAGTCTTGGGCGCTGGAGGGCCTGTCCTGACC**ATG**  
 GTCCCTGCCTGGCTGTGGCTGCTTTGTGTCTCCGTCCCCCAGGCTCTCCCCAAGGCCAGCCTGC  
 AGAGCTGTCTGTGGAAGTTCAGAAAACTATGGTGGAAATTTCCCTTTATACCTGACCAAGTTGC  
 CGCTGCCCCGTGAGGGGGCTGAAGGCCAGATCGTGCTGTCAGGGGACTCAGGCAAGGCAACTGAG  
 GGCCCATTTGCTATGGATCCAGATTCTGGCTTCCTGCTGGTGACCAGGGCCCTGGACCGAGAGGA  
 GCAGGCAGAGTACCAGCTACAGGTCACCCTGGAGATGCAGGATGGACATGTCTTGTGGGGTCCAC  
 AGCCTGTGCTTGTGCACGTGAAGGATGAGAATGACCAGGTGCCCCATTTCTCTCAAGCCATCTAC  
 AGAGCTCGGCTGAGCCGGGGTACCAGGCCTGGCATCCCCTTCTCTTCCCTTGAGGCTTCAGACCG  
 GGATGAGCCAGGCACAGCCAACTCGGATCTTCGATTCCACATCCTGAGCCAGGCTCCAGCCCAGC  
 CTTCCCCAGACATGTTCCAGCTGGAGCCTCGGCTGGGGGCTCTGGCCCTCAGCCCCAAGGGGAGC  
 ACCAGCCTTGACCACGCCCTGGAGAGGACCTACCAGCTGTTGGTACAGGTCAAGGACATGGGTGA  
 CCAGGCCTCAGGCCACCAGGCCACTGCCACCGTGGAAGTCTCCATCATAGAGAGCACCTGGGTGT  
 CCCTAGAGCCTATCCACCTGGCAGAGAATCTCAAAGTCCTATACCCGCACCACATGGCCCAGGTA  
 CACTGGAGTGGGGGTGATGTGCACTATCACCTGGAGAGCCATCCCCGGGACCCTTTGAAGTGAA  
 TGCAGAGGGAAACCTCTACGTGACCAGAGAGCTGGACAGAGAAGCCCAGGCTGAGTACCTGCTCC  
 AGGTGCGGGCTCAGAATTCCCATGGCGAGGACTATGCGGCCCTCTGGAGCTGCACGTGCTGGTG  
 ATGGATGAGAATGACAACGTGCCATCTGCCCTCCCCGTGACCCACAGTCAGCATCCCTGAGCT  
 CAGTCCACCAGGTACTGAAGTGACTAGACTGTCAGCAGAGGATGCAGATGCCCCCGGCTCCCCCA  
 ATTCCCACGTTGTGTATCAGCTCCTGAGCCCTGAGCCTGAGGATGGGGTAGAGGGGAGAGCCTTC  
 CAGGTGGACCCCACTTCAGGCAGTGTGACGCTGGGGGTGCTCCCACTCCGAGCAGGCCAGAACAT  
 CCTGCTTCTGGTGCTGGCCATGGACCTGGCAGGCGCAGAGGGTGGCTTCAGCAGCACGTGTGAAG  
 TCGAAGTCGCAGTCACAGATATCAATGATCACGCCCTGAGTTCATCACTTCCCAGATTGGGCCT  
 ATAAGCCTCCCTGAGGATGTGGAGCCCGGACTCTGGTGGCCATGCTAACAGCCATTGATGCTGA  
 CCTCGAGCCCGCCTTCCGCCTCATGGATTTTGCCATTGAGAGGGGAGACACAGAAGGGACTTTTG  
 GCCTGGATTGGGAGCCAGACTCTGGGCATGTTAGACTCAGACTCTGCAAGAACCTCAGTTATGAG  
 GCAGCTCCAAGTCATGAGGTGGTGGTGGTGGTGAGAGTGTGGCGAAGCTGGTGGGGCCAGGCCC  
 AGGCCCTGGAGCCACCGCCACGGTGACTGTGCTAGTGAGAGAGTGTGATGCCACCCCCCAAGTTGG  
 ACCAGGAGAGCTACGAGGCCAGTGTCCCCATCAGTGCCCCAGCCGGCTCTTTCCTGCTGACCATC  
 CAGCCCTCCGACCCCATCAGCCGAACCTCAGGTTCTCCCTAGTCAATGACTCAGAGGGCTGGCT  
 CTGCATTGAGAAATTCTCCGGGGAGGTGCACACCGCCAGTCCCTGCAGGGCGCCCAGCCTGGGG  
 ACACCTACACGGTGCTTGTGGAGGCCAGGATACAGCCCTGACTCTTGCCCCTGTGCCCTCCCAA  
 TACCTCTGCACACCCCGCCAAGACCATGGCTTGATCGTGAGTGGACCCAGCAAGGACCCCGATCT  
 GGCCAGTGGGCACGGTCCCTACAGCTTCACCCTTGGTCCCAACCCACGGTGCAACGGGATTGGC  
 GCCTCCAGACTCTCAATGGTTCCCATGCCTACCTCACCTTGGCCCTGCATTGGGTGGAGCCACGT  
 GAACACATAATCCCCGTGGTGGTCAGCCACAATGCCAGATGTGGCAGCTCCTGGTTTCGAGTGAT  
 CGTGTGTGCTGCAACGTGGAGGGGCAGTGCATGCGCAAGGTGGGCCGATGAAGGGCATGCCCA  
 CGAAGCTGTCGGCAGTGGGCATCCTTGTTAGGCACCCTGGTAGCAATAGGAATCTTCCTCATCCTC  
 ATTTTACCCACTGGACCATGTCAAGGAAGAAGGACCCGGATCAACCAGCAGACAGCGTGCCCCCT  
 GAAGGCGACTGTCT**TGA**ATGGCCCAGGCAGCTCTAGCTGGGAGCTTGGCCTCTGGCTCCATCTGAG  
 TCCCCTGGGAGAGAGCCCAGCACCCAAAGATCCAGCAGGGGACAGGACAGAGTAGAAGCCCCCTCCA  
 TCTGCCCTGGGGTGGAGGCACCATCACCATCACAGGCATGTCTGCAGAGCCTGGACACCAACTT  
 TATGGACTGCCCATGGGAGTGCTCCAAATGTCAGGGTGTTTGCCCAATAATAAAGCCCCAGAGAA  
 CTGGGCTGGGCCCTATGGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAG

**FIGURE 98**

MVPAWLWLLCVSVPQALPKAQPAELSVPEPENYGGNFPLYLTKLPLPREGAEGQIVLSGDSGKAT  
EGPFAMDPDSGFLLVTRALDREEQAEYQLQVTLEMQDGHVLWGPQPVLVHVKDENDQVPHFSQAI  
YRARLSRGTRPGIPFLFLEASDRDEPGTANSDLRFHILSQAPAQSPDMFQLEPRLGALALSPKG  
STSLDHALERTYQLLVQVKDMGDQASGHQATATVEVSIIESTWVSLEPIHLAENLKVLYPHHMAQ  
VHWSGGDVHYHLESHPPGPFEVNAEGNLYVTRELDREAQAEYLLQVRAQNSHGEDIAAPLELHVL  
VMDENDNVPICPPRDPTVSIPELSPPGTEVTRLAEDADAPGSPNSHVYQLLSPEPEDGVEGRA  
FQVDPTSGSVTLGVLPLRAGQNILLVLAMDLAGAEGGFSSTCEVEVAVTDINDHAPEFITSQIG  
PISLPEDVEPGTLVAMLTADLEPAFRLMDFAIERGDTEGTFGLDWEPPDSGHVRLRLCKNLSY  
EAAPSHEVVVVQSVAKLVGPGPGGATATVTVLVERVMPPPKLDQESYEASVPI SAPAGSFLLT  
IQPSDPISRTLRFSLVNDSEGWLCIEKFSGEVHTAQSLQGAQPGDTYTVLVEAQDTALT LAPVPS  
QYLCTPRQDHGLIVSGPSKDPDLASGHGPYSFTLGNPTVQRDWRLQTLNGSHAYLTLALHWVEP  
REHIIPVVVSHNAQMWQLLVRVIVCRCNVEGQCMRKVGRMKGMPTKLSAVGILVGTLVAIGIFLI  
LIFTHWTMSRKKDPDQPADSVPLKATV

**Signal peptide:**

amino acids 1-18

**Transmembrane domain:**

amino acids 762-784

**FIGURE 99**

GGCTGACCGTGCTACATTGCCTGGAGGAAGCCTAAGGAACCCAGGCATCCAGCTGCCCACGCCTG  
AGTCCAAGATTCTTCCCAGGAACACAAACGTAGGAGACCCACGCTCCTGGAAGCACCAGCCTTTA  
TCTCTTCACCTTCAAGTCCCCTTTCTCAAGAATCCTCTGTTCTTTGCCCTCTAAAGTCTTGGTAC  
ATCTAGGACCCAGGCATCTTGCTTTCCAGCCACAAAGAGACAGATGAAGATGCAGAAAGGAAATG  
TTCTCCTTATGTTTGGTCTACTATTGCATTTAGAAGCTGCAACAAATTCCAATGAGACTAGCACC  
TCTGCCAACACTGGATCCAGTGTGATCTCCAGTGGAGCCAGCACAGCCACCAACTCTGGGTCCAG  
TGTGACCTCCAGTGGGGTCAGCACAGCCACCATCTCAGGGTCCAGCGTGACCTCCAATGGGGTCA  
GCATAGTCACCAACTCTGAGTTCATACAACCTCCAGTGGGATCAGCACAGCCACCAACTCTGAG  
TTCAGCACAGCGTCCAGTGGGATCAGCATAGCCACCAACTCTGAGTCCAGCACAACTCCAGTGG  
GGCCAGCACAGCCACCAACTCTGAGTCCAGCACACCCTCCAGTGGGGCCAGCACAGTCACCAACT  
CTGGGTCCAGTGTGACCTCCAGTGGAGCCAGCACTGCCACCAACTCTGAGTCCAGCACAGTGTCC  
AGTAGGGCCAGCACTGCCACCAACTCTGAGTCTAGCACACTCTCCAGTGGGGCCAGCACAGCCAC  
CAACTCTGACTCCAGCACAACTCCAGTGGGGCTAGCACAGCCACCAACTCTGAGTCCAGCACAA  
CCTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACAGTGTCCAGTAGGGCCAGCACT  
GCCACCAACTCTGAGTCCAGCACAACTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAG  
AACGACCTCCAATGGGGCTGGCACAGCCACCAACTCTGAGTCCAGCACGACCTCCAGTGGGGCCA  
GCACAGCCACCAACTCTGACTCCAGCACAGTGTCCAGTGGGGCCAGCACTGCCACCAACTCTGAG  
TCCAGCACGACCTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACGACCTCCAGTGG  
GGCTAGCACAGCCACCAACTCTGACTCCAGCACAACTCCAGTGGGGCCGGCACAGCCACCAACT  
CTGAGTCCAGCACAGTGTCCAGTGGGATCAGCACAGTCACCAATTCTGAGTCCAGCACACCCTCC  
AGTGGGGCCAACACAGCCACCAACTCTGAGTCCAGTACGACCTCCAGTGGGGCCAACACAGCCAC  
CAACTCTGAGTCCAGCACAGTGTCCAGTGGGGCCAGCACTGCCACCAACTCTGAGTCCAGCACAA  
CCTCCAGTGGGGTCAGCACAGCCACCAACTCTGAGTCCAGCACAACTCCAGTGGGGCTAGCACA  
GCCACCAACTCTGACTCCAGCACAACTCCAGTGAGGCCAGCACAGCCACCAACTCTGAGTCTAG  
CACAGTGTCCAGTGGGATCAGCACAGTCACCAATTCTGAGTCCAGCACAACTCCAGTGGGGCCA  
ACACAGCCACCAACTCTGGGTCCAGTGTGACCTCTGCAGGCTCTGGAACAGCAGCTCTGACTGGA  
ATGCACACAACTTCCCATAGTGCATCTACTGCAGTGAGTGAGGCAAAGCCTGGTGGGTCCCTGGT  
GCCGTGGGAAATCTTCCTCATCACCTGGTCTCGGTTGTGGCGGCCGTGGGGCTCTTTGCTGGGC  
TCTTCTTCTGTGTGAGAAACAGCCTGTCCCTGAGAAACACCTTTAACACAGCTGTCTACCACCT  
CATGGCCTCAACCATGGCCTTGGTCCAGGCCCTGGAGGGAATCATGGAGCCCCCACAGGCCAG  
GTGGAGTCTTAACCTGGTTCTGGAGGAGACCAGTATCATCGATAGCCATGGAGATGAGCGGGAGGA  
ACAGCGGGCCCTGAAGCAGCCCCGGAAGCAAGTGCCGCATTCTTCAGGAAGGAAGAGACCTGGGCA  
CCCAAGACCTGGTTTCCTTTTCATTCATCCCAGGAGACCCCTCCCAGCTTTGTTTGAGATCCTGAA  
AATCTTGAAGAAGGTATTCCCTCACCTTTCTTGCCTTTACCAGACACTGGAAAGAGAATACTATAT  
TGCTCATTTAGCTAAGAAATAAATACATCTCATCTAACACACACGACAAAGAGAAGCTGTGCTTG  
CCCCGGGGTGGGTATCTAGCTCTGAGATGAACTCAGTTATAGGAGAAAACCTCCATGCTGGACTC  
CATCTGGCATTCAAAATCTCCACAGTAAATCCAAAGACCTCAAAAAAAAAAAAAAAAAAAAAA  
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

## **FIGURE 100**

MKMQKGNVLLMFGLLLHLEAATNSNETSTSANTGSSVISSGASTATNSGSSVTSSGVSTATISGS  
SVTSNGVSIVTNSEFHTTSSGISTATNSEFSTASSGISIATNSESSTTSSGASTATNSESSTPSS  
GASTVTNSGSSVTSSGASTATNSESSTVSSRASTATNSESSTLSSGASTATNSDSSTTSSGASTA  
TNSESSTTSSGASTATNSESSTVSSRASTATNSESSTTSSGASTATNSESRTTSNGAGTATNSES  
STTSSGASTATNSDSSTVSSGASTATNSESSTTSSGASTATNSESSTTSSGASTATNSDSSTTSS  
GAGTATNSESSTVSSGISTVTNSESSTPSSGANTATNSESSTTSSGANTATNSESSTVSSGASTA  
TNSESSTTSSGVSTATNSESSTTSSGASTATNSDSSTTSSEASTATNSESSTVSSGISTVTNSES  
STTSSGANTATNSGSSVTSAGSGTAALTGMHTTSHSASTAVSEAKPGGSLVPWEIFLITLVSVVA  
AVGLFAGLFFCVRNSLSLRNTFNTAVYHPHGLNHGLGPGPGGNHGAPHRPRWSPNWFRRPVSSI  
AMEMSGRNSGP

**Signal peptide:**

amino acids 1-20

**Transmembrane domain:**

amino acids 510-532

**FIGURE 101**

GGCCGGACGCCTCCGCGTTACGGGATGAATTAACGGCGGGTTCCGCACGGAGGTTGTGACCCCTA  
CGGAGCCCCAGCTTGCCACGCACCCCACTCGGCGTCGCGCGGCGTGCCCTGCTTGTCACAGGTG  
GGAGGCTGGAACATCAGGCTGAAAAACAGAGTGGGTACTCTCTTCTGGGAAGCTGGCAACAAAT  
GGATGATGTGATATATGCATTCCAGGGGAAGGGAAATTGTGGTGCTTCTGAACCCATGGTCAATT  
AACGAGGCAGTTTCTAGCTACTGCACGTACTTCATAAAGCAGGACTCTAAAAGCTTTGGAATCAT  
GGTGTCATGGAAAGGGATTTACTTTATACTGACTCTGTTTTGGGGAAGCTTTTTTGGGAAGCATTT  
TCATGCTGAGTCCCTTTTTACCTTTGATGTTTGTAACCCATCTTGGTATCGCTGGATCAACAAC  
CGCCTTGTTGGCAACATGGCTCACCTACCTGTGGCATTATTGGAGACCATGTTTGGTGTAAGT  
GATTATAACTGGGGATGCATTTGTTCTTGAGAAAGAAGTGTCAATTATCATGAACCATCGGACAA  
GAATGGACTGGATGTTCTGTGGAATTGCCTGATGCGATATAGCTACCTCAGATTGGAGAAAATT  
TGCTCAAAGCGAGTCTCAAAGGTGTTCTGGATTTGGTTGGGCCATGCAGGCTGCTGCCTATAT  
CTTCATTCATAGGAAATGGAAGGATGACAAGAGCCATTTCTGAAGACATGATTGATTACTTTTGTG  
ATATTCACGAACCACTTCAACTCCTCATATTTCCAGAAGGGACTGATCTCACAGAAAACAGCAAG  
TCTCGAAGTAATGCATTTGCTGAAAAAATGGACTTCAGAAATATGAATATGTTTTACATCCAAG  
AACTACAGGCTTTACTTTTGTGGTAGACCGTCTAAGAGAAGGTAAGAACCCTTGATGCTGTCCATG  
ATATCACTGTGGCGTATCCTCACAACATTCCTCAATCAGAGAAGCACCTCCTCCAAGGAGACTTT  
CCCAGGGAAATCCACTTTTACGTCCACCGGTATCCAATAGACACCCTCCCCACATCCAAGGAGGA  
CCTTCAACTCTGGTGCCACAAACGGTGGGAAGAGAAAGAAGAGAGGCTGCGTTCCTTCTATCAAG  
GGGAGAAGAATTTTTATTTTACCGGACAGAGTGTCAATTCCACCTTGCAAGTCTGAACTCAGGGTC  
CTTGTGGTCAAATTGCTCTCTATACTGTATTGGACCCTGTTAGCCCTGCAATGTGCCTACTCAT  
ATATTTGTACAGTCTTGTTAAGTGGTATTTTATAATCACCATTGTAATCTTTGTGCTGCAAGAGA  
GAATATTTGGTGGACTGGAGATCATAGAACTGCATGTTACCGACTTTTACACAAACAGCCACAT  
TTAAATTCAAAGAAAAATGAGTAAGATTATAAGGTTTGCCATGTGAAAACCTAGAGCATATTTTG  
GAAATGTTCTAAACCTTTCTAAGCTCAGATGCATTTTTGCATGACTATGTGAATATTTCTTACT  
GCCATCATTATTTGTAAAGATATTTTGCACCTAATTTTGTGGGAAAAATATTGCTACAATTTTT  
TTAATCTCTGAATGTAATTTGATACTGTGTACATAGCAGGGAGTGATCGGGGTGAAATAACTT  
GGGCCAGAATATTATTAAACAATCATCAGGCTTTTAA

## **FIGURE 102**

MHSRGREIVVLLNPWSINEAVSSYCTYFIKQDSKSGIMVSWKGIYFILTLFWGSFFGSI FMLS P  
FLPLMFVNPSWYRWINNRLVATWLTLPVALLETMFGVKVIITGDAFVPGERSVIIMNHRTRMDWM  
FLWNCLMRYSYLRLEKICLKASLKGVPFGFWAMQAAAYIFIHRKWKDDKSHFEDMIDYFCDIHEP  
LQLLIFPEGTDLTENSKSRSNFAEKNGLQKYEYVLHPRTTGFTFVVDRLREGKNLDAVHDITVA  
YPHNIPQSEKHL LQGDFPREIH FHVHRYPIDTLPTS KEDLQLWCHKRWEEKEERLRSFYQGEKNF  
YFTGQSVIPPCKSELRVLVVKLLSILYWTLFSPAMCLLIYLYSLVKWYFIITIVIFVLQERIFGG  
LEIIE LACYRLLHKQPHLNSKKNE

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-22

#### **Transmembrane domains:**

amino acids 44-63, 90-108, 354-377



**FIGURE 103**

CGGCTCGAGCGGCTCGAGTGAAGAGCCTCTCCACGGCTCCTGCGCCTGAGACAGCTGGCCTGACC  
 TCCAAATCATCCATCCACCCCTGCTGTCATCTGTTTTCATAGTGTGAGATCAACCCACAGGAATA  
 TCC**ATG**GCTTTTGTGCTCATTTTGGTTCTCAGTTTCTACGAGCTGGTGTGAGGACAGTGGCAAGT  
 CACTGGACCGGGCAAGTTTGTCCAGGCCTTGGTGGGGGAGGACGCCGTGTTCTCCTGCTCCCTCT  
 TTCCTGAGACCAGTGCAGAGGCTATGGAAGTGC GGTTCTTCAGGAATCAGTTCCATGCTGTGGTC  
 CACCTCTACAGAGATGGGGAAGACTGGGAATCTAAGCAGATGCCACAGTATCGAGGGGAGAACTGA  
 GTTGTGAAGGACTCCATTGCAGGGGGGCGTGTCTCTCTAAGGCTAAAAACATCACTCCCTCGG  
 ACATCGGCCTGTATGGGTGCTGGTTCAGTTCCCAGATTTACGATGAGGAGGCCACCTGGGAGCTG  
 CGGGTGGCAGCACTGGGCTCACTTCCTCTCATTTCCATCGTGGGATATGTTGACGGAGGTATCCA  
 GTTACTCTGCCTGTCCTCAGGCTGGTTCCCCCAGCCACAGCCAAGTGGAAGGTCCACAAGGAC  
 AGGATTTGTCTTCAGACTCCAGAGCAAATGCAGATGGGTACAGCCTGTATGATGTGGAGATCTCC  
 ATTATAGTCCAGGAAAATGCTGGGAGCATATTGTGTTCCATCCACCTTGCTGAGCAGAGTCATGA  
 GGTGGAATCCAAGGTATTGATAGGAGAGACGTTTTTCCAGCCCTCACCTTGCGCCTGGCTTCTA  
 TTTTACTCGGGTTACTCTGTGGTGCCCTGTGTGGTGTTCATGGGGATGATAATTGTTTTCTTC  
 AAATCCAAAGGGAAAATCCAGGCGGAAGTGGACTGGAGAAGAAAGCACGGACAGGCAGAAATTGAG  
 AGACGCCCCGAAACACGCAGTGGAGGTGACTCTGGATCCAGAGACGGCTCACCCGAAGCTCTGCG  
 TTTCTGATCTGAAAAGTGTAAACCATAGAAAAGCTCCCCAGGAGGTGCCTCACTCTGAGAAGAGA  
 TTTACAAGGAAGAGTGTGGTGGCTTCTCAGGGTTTCCAAGCAGGGAGACATTACTGGGAGGTGGA  
 CGTGGGACAAAATGTAGGGTGGTATGTGGGAGTGTGTCGGGATGACGTAGACAGGGGGAAGAACA  
 ATGTGACTTTGTCTCCCAACAATGGGTATTGGGTCCCTCAGACTGACAACAGAACATTTGTATTTTC  
 ACATTCAATCCCCATTTTATCAGCCTCCCCCCCCAGCACCCCTCCTACACGAGTAGGGGTCTTCCT  
 GGACTATGAGGGTGGGACCATCTCCTTCTTCAATACAAATGACCAGTCCCTTATTTATACCCTGC  
 TGACATGTCAGTTTGAAGGCTTGTGAGACCCTATATCCAGCATGCGATGTATGACGAGGAAAAG  
 GGGACTCCCATATTATATGTCCAGTGTCTTGGGGAT**TGA**GACAGAGAAGACCCTGCTTAAAGGGC  
 CCCACACCACAGACCCAGACACAGCCAAGGGAGAGTGCTCCCGACAGGTGGCCCCAGCTTCCTCT  
 CCGGAGCCTGCGCACAGAGAGTCACGCCCCCACTCTCCTTTAGGGAGCTGAGGTTCTTCTGCCC  
 TGAGCCCTGCAGCAGCGGCAGTCACAGCTTCCAGATGAGGGGGGATTGGCCTGACCCTGTGGGAG  
 TCAGAAGCCATGGCTGCCCTGAAGTGGGGACGGAATAGACTCACATTAGGTTTAGTTTGTGAAAA  
 CTCATCCAGCTAAGCGATCTTGAACAAGTCACAACCTCCCAGGCTCCTCATTTGCTAGTCACGG  
 ACAGTGATTCCCTGCCTCACAGGTGAAGATTAAAGAGACAACGAATGTGAATCATGCTTGCAGGTT  
 TGAGGGCACAGTGTGTGCTAATGATGTGTTTTTATATTATACATTTTCCCACCATAAACTCTGTT  
 TGCTTATTCCACATTAATTTACTTTTCTCTATACCAAATCACCCATGGAATAGTTATTGAACACC  
 TGCTTTGTGAGGCTCAAAGAATAAAGAGGAGGTAGGATTTTTCACTGATTCTATAAGCCCAGCAT  
 TACCTGATACCAAAACCAGGCAAAGAAAACAGAAGAAGAGGAAGGAAAACCTACAGGTCCATATCC  
 CTCATTAACACAGACACAAAAATTCTAAATAAAATTTTAACAAATTAACTAAACAATATATTTA  
 AAGATGATATATACTACTCAGTGTGGTTTGTCCCACAAATGCAGAGTTGGTTTAATATTTAAAT  
 ATCAACCAGTGTAATTCAGCACATTAATAAAGTAAAAAAGAAAACCATAAAAAAAAAAAAAAA

## **FIGURE 104**

MAFVLILVLSFYELVSGQWQVTGPGKFVQALVGEDAVFSCSLFPETSAEAMEVRFFRNQFHAVVH  
LYRDGEDWESKQMPQYRGRTEFVKDSIAGGRVSLRLKNITPSDIGLYGCWFSSQIYDEEATWELR  
VAALGSLPLISIVGYVDGGIQLLCLSSGWFPQPTAKWKGPQGQDLSSDSRANADGYSLYDVEISI  
IVQENAGSILCSIHLAEQSHEVESKVLIGETFFQPSPWRLASILLGLLCGALCGVVMGMIIVFFK  
SKGKIQAELDWRRKHGQAELRDARKHAVEVTLDPETAHPKLCVSDLKTVTHRKAPQEVPHSEKRF  
TRKSVVASQGFQAGRHYWEVDVGQNVGWYVGVCRDDVDRGKNNVTLSPNNGYWVLRLTTEHLYFT  
FNPHFISLPPSTPPTRVGVFLDYEGGTISFFNTNDQS LIYTL LTCQFEGLLRPYIQHAMYDEEKG  
TPIFICPVSWG

**Signal peptide:**

amino acids 1-17

**Transmembrane domains:**

amino acids 131-150, 235-259

**FIGURE 105**

CCTTCACAGGACTCTTCATTGCTGGTTGGCA**ATG**ATGTATCGGCCAGATGTGGTGAGGGCTAGGA  
AAAGAGTTTGTGTTGGGAACCCTGGGTTATCGGCCCTCGTCATCTTCATATCCCTGATTGTCCTGGCA  
GTGTGCATTGGACTCACTGTTTCATTATGTGAGATATAATCAAAAGAAGACCTACAATTACTATAG  
CACATTGTCATTTACAACGACAACTATATGCTGAGTTTGGCAGAGAGGCTTCTAACAATTTTA  
CAGAAATGAGCCAGAGACTTGAATCAATGGTGAAAAATGCATTTTATAAATCTCCATTAAGGGAA  
GAATTTGTCAAGTCTCAGGTTATCAAGTTCAGTCAACAGAAGCATGGAGTGTGAGTGCATATGCT  
GTTGATTTGTAGATTTCACTCTACTGAGGATCCTGAAACTGTAGATAAAATTGTTCAACTTGTTT  
TACATGAAAAGCTGCAAGATGCTGTAGGACCCCTAAAGTAGATCCTCACTCAGTTAAAATTAAA  
AAAATCAACAAGACAGAAACAGACAGCTATCTAAACCATTGCTGCGGAACACGAAGAAGTAAAAC  
TCTAGGTGAGAGTCTCAGGATCGTTGGTGGGACAGAAGTAGAAGAGGGTGAATGGCCCTGGCAGG  
CTAGCCTGCAGTGGGATGGGAGTCATCGCTGTGGAGCAACCTTAATTAATGCCACATGGCTTGTG  
AGTGCTGCTCACTGTTTTACAACATATAAGAACCCTGCCAGATGGACTGCTTCCTTTGGAGTAAC  
AATAAACCTTCGAAAATGAAACGGGGTCTCCGGAGAATAATTGTCCATGAAAAATACAAACACC  
CATCACATGACTATGATATTTCTCTTGACAGAGCTTTCTAGCCCTGTTCCCTACACAAATGCAGTA  
CATAGAGTTTGTCTCCCTGATGCATCCTATGAGTTTCAACCAGGTGATGTGATGTTTGTGACAGG  
ATTTGGAGCACTGAAAAATGATGGTTACAGTCAAAATCATCTTCGACAAGCACAGGTGACTCTCA  
TAGACGCTACAACCTGCAATGAACCTCAAGCTTACAATGACGCCATAACTCCTAGAATGTTATGT  
GCTGGCTCCTTAGAAGGAAAAACAGATGCATGCCAGGGTGACTCTGGAGGACCACTGGTTAGTTC  
AGATGCTAGAGATATCTGGTACCTTGCTGGAATAGTGAGCTGGGGAGATGAATGTGCGAAACCCA  
ACAAGCCTGGTGTTTATACTAGAGTTACGGCCTTGCGGGACTGGATTACTTCAAAAACCTGGTATC  
**TAAG**AGACAAAAGCCTCATGGAACAGATAACATTTTTTTTTTTGTTTTTTGGGTGTGGAGGCCATTT  
TTAGAGATACAGAATTGGAGAAGACTTGCAAAACAGCTAGATTTGACTGATCTCAATAAACTGTT  
TGCTTGATGCATGTATTTTCTTCCCAGCTCTGTTCCGCACGTAAGCATCCTGCTTCTGCCAGATC  
AACTCTGTCATCTGTGAGCAATAGTTGAAACTTTATGTACATAGAGAAATAGATAATACAATATT  
ACATTACAGCCTGTATTCATTTGTTCTCTAGAAGTTTTTGTGAGAATTTTGACTTGTTGACATAAA  
TTTGTAATGCATATATACAATTTGAAGCACTCCTTTTCTTCAGTTCCTCAGCTCCTCTCATTTCA  
GCAATATCCATTTTCAAGGTGCAGAACAAAGGAGTGAAAGAAAATATAAGAAGAAAAAAATCCCC  
TACATTTTATTGGCACAGAAAAGTATTAGGTGTTTTTCTTAGTGGAATATTAGAAATGATCATAT  
TCATTATGAAAGGTCAAGCAAAGACAGCAGAATACCAATCACTTCATCATTTAGGAAGTATGGGA  
ACTAAGTTAAGGAAGTCCAGAAAGAAGCCAAGATATATCCTTATTTTCATTTCCAAACAACACTACT  
ATGATAAATGTGAAGAAGATTCTGTTTTTTGTGACCTATAATAATTATACAACTTCATGCAAT  
GTACTTGTTCTAAGCAAATTAAAGCAAATATTTATTTAACATTGTTACTGAGGATGTCAACATAT  
AACATAAAATATAAATCACCCA

## **FIGURE 106**

MMYRPDVVRARKRVCWEPWVIGLVIFISLIVLAVCIGLTVHYVRYNQKKTYNYYSTLSFTTDKLY  
AEFGREASNNFTEMSQRLESMVKNAFYKSPLREEFVKSQVIKFSQQKHGVLAHMLLICRFHSTED  
PETVDKIVQLVLHEKLQDAVGPPKVDPHSVKIKKINKTETDSYLNHCCGTRRSKTLGQSLRIVGG  
TEVEEGEWPWQASLQWDGSHRCGATLINATWLVSAAHCFTTYKNPARWTASFGVTIKPSKMKRGL  
RRIIVHEKYKHPSHDYDISLAELSSPVPTYNAVHRVCLPDASYEFQPGDVMFVTGFGALKNDGYS  
QNHLRQAQVTLIDATTCNEPQAYNDAITPRMLCAGSLEGKTDACQGDSGGPLVSSDARDIWYLAG  
IVSWGDECAKPNKPGVYTRVTALRDWITSKTGI

**Transmembrane domain:**

amino acids 21-40 (type II)

**FIGURE 107**

AGAGAAAGAAGCGTCTCCAGCTGAAGCCAATGCAGCCCTCCGGCTCTCCGCGAAGAAGTTCCCTG  
CCCCGATGAGCCCCCGCCGTGCGTCCCCGACTATCCCCAGGCGGGCGTGGGGCACCGGGCCCAGC  
GCCGACGATCGCTGCCGTTTTGCCCTTGGGAGTAGGATGTGGTGAAAGGATGGGGCTTCTCCCTT  
ACGGGGCTCACAA**ATG**GCCAGAGAAGATTCCGTGAAGTGTCTGCGCTGCCTGCTCTACGCCCTCAA  
TCTGCTCTTTTGGTTAATGTCCATCAGTGTGTTGGCAGTTTCTGCTTGGATGAGGGACTACCTAA  
ATAATGTTCTCACTTTAACTGCAGAAACGAGGGTAGAGGAAGCAGTCATTTTGACTTACTTTTCCT  
GTGGTTTCATCCGGTCATGATTGCTGTTTGCTGTTTCCTTATCATTGTGGGGATGTTAGGATATTG  
TGGAACGGTGAAAAGAAATCTGTTGCTTCTTGCATGGTACTTTGGAAGTTTGCTTGTCAATTTTCT  
GTGTAGAACTGGCTTGTGGCGTTTGGACATATGAACAGGAACCTTATGGTTCCAGTACAATGGTCA  
GATATGGTCACTTTGAAAGCCAGGATGACAAATTATGGATTACCTAGATATCGGTGGCTTACTCA  
TGCTTGGAATTTTTTTTCAGAGAGAGTTTAAGTGCTGTGGAGTAGTATATTTCACTGACTGGTTGG  
AAATGACAGAGATGGACTGGCCCCCAGATTCTGCTGTGTTAGAGAATTTCCAGGATGTTCCAAA  
CAGGCCCCACCAGGAAGATCTCAGTGACCTTTATCAAGAGGGTGTGGGAAGAAAATGTATTCCTT  
TTTGAGAGGAACCAAACAACCTGCAGGTGCTGAGGTTTCTGGGAATCTCCATTGGGGTGACACAAA  
TCCTGGCCATGATTCTCACCATTACTCTGCTCTGGGCTCTGTATTATGATAGAAGGGAGCCTGGG  
ACAGACCAAATGATGTCCTTGAAGAATGACAACTCTCAGCACCTGTCATGTCCCTCAGTAGAACT  
GTTGAAACCAAGCCTGTCAAGAATCTTTGAACACACATCCATGGCAAACAGCTTTAATACACACT  
TTGAGATGGAGGAGTTA**TAA**AAAGAAATGTCACAGAAGAAAACCACAACTTGTTTTATTGGACT  
TGTGAATTTTTGAGTACATACTATGTGTTTCAGAAATATGTAGAAATAAAAATGTTGCCATAAAA  
TAACACCTAAGCATATACTATTCTATGCTTTAAAATGAGGATGGAAAAGTTTCATGTCATAAGTC  
ACCACCTGGACAATAATTGATGCCCTTAAAATGCTGAAGACAGATGTCATACCCACTGTGTAGCC  
TGTGTATGACTTTTTACTGAACACAGTTATGTTTTGAGGCAGCATGGTTTGATTAGCATTTCCGCA  
TCCATGCAAACGAGTCACATATGGTGGGACTGGAGCCATAGTAAAGGTTGATTACTTCTACCAA  
CTAGTATATAAAGTACTAATTAAATGCTAACATAGGAAGTTAGAAAATACTAATAACTTTTATTA  
CTCAGCGATCTATTCTTCTGATGCTAAATAAATTATATATCAGAAAACCTTCAATATTGGTGACT  
ACCTAAATGTGATTTTTGCTGGTTACTAAAATATTCTTACCCTTAAAAGAGCAAGCTAACACAT  
TGTCTTAAGCTGATCAGGGATTTTTTGTATATAAGTCTGTGTTAAATCTGTATAATTCAGTCGAT  
TTCAGTTCTGATAATGTTAAGAATAACCATTATGAAAAGGAAAATTTGTCCTGTATAGCATCATT  
ATTTTTAGCCTTTTCCTGTTAATAAAGCTTTACTATTCTGTCCTGGGCTTATATTACACATATAAC  
TGTTATTTAAATACTTAACCACTAATTTTGAAAATTACCAGTGTGATACATAGGAATCATTATTC  
AGAATGTAGTCTGGTCTTTAGGAAGTATTAATAAGAAAATTTGCACATAACTTAGTTGATTCAGA  
AAGGACTTGATGCTGTTTTTCTCCCAAATGAAGACTCTTTTTTGACACTAAACACTTTTTAAAAA  
GCTTATCTTTGCCTTCTCCAAACAAGAAGCAATAGTCTCCAAGTCAATATAAATTTCTACAGAAAA  
TAGTGTTCTTTTTCTCCAGAAAAATGCTTGTGAGAATCATTAACACATGTGACAAATTTAGAGATT  
CTTTGTTTTATTTCACTGATTAATATACTGTGGCAAATTACACAGATTATTAAATTTTTTTTACAA  
GAGTATAGTATATTTATTTGAAATGGGAAAAGTGCATTTTACTGTATTTTGTGTATTTTGTGTTAT  
TTCTCAGAATATGGAAAGAAAATTAATAATGTGTCAATAAATATTTTCTAGAGAGTAA

## **FIGURE 108**

MAREDSVKCLRCLLYALNLLFWLMSISVLAVSAWMRDYLNNVLTTLTAETRVEEAVILTYFPVVHP  
VMIAVCCFLIIVGMLGYCGTVKRNLLLLAWYFGSLLVIFCVELACGVWTYEQELMVPVQWSDMVT  
LKARMTNYGLPRYRWLTHAWNFFQREFKCCGVVYFTDWLEMTEMDWPPDSCCVREFPGCSKQAHQ  
EDLSDLYQEGCGKKMYSFLRGTKQLQVLRFLGISIGVTQILAMILTITLLWALYYDRREP GTDQM  
MSLKNDNSQHLSCPSVELLKPSLSRIFEHTSMANSFNTHFEMEEL

**Signal peptide:**

amino acids 1-33

**Transmembrane domains:**

amino acids 12-35, 57-86, 94-114, 226-248

**FIGURE 109**

CCAAGGCCAGAGCTGTGGACACCTTATCCCACTCATCCTCATCCTCTTCCTCTGATAAAGCCCCT  
ACCAAGTGCTGATAAAGTCTTTCTCGTGAGAGCCTAGAGGCCTTAAAAAAAAAAGTGCTTGAAAGA  
GAAGGGGACAAAGGAACACCAGTATTAAGAGGATTTTCCAGTGTTTCTGGCAGTTGGTCCAGAAG  
**GATG**CCTCCATTCCCTGCTTCTCACCTGCCTCTTCATCACAGGCACCTCCGTGTCACCCGTGGCCC  
TAGATCCTTGTTCTGCTTACATCAGCCTGAATGAGCCCTGGAGGAACACTGACCACCAGTTGGAT  
GAGTCTCAAGGTCCTCCTCTATGTGACAACCATGTGAATGGGGAGTGGTACCACTTCACGGGCAT  
GGCGGGAGATGCCATGCCTACCTTCTGCATACCAGAAAACCACTGTGGAACCCACGCACCTGTCT  
GGCTCAATGGCAGCCACCCCCTAGAAGGCGACGGCATTGTGCAACGCCAGGCTTGTGCCAGCTTC  
AATGGGAACTGCTGTCTCTGGAACACCACGGTGGAAGTCAAGGCTTGCCCTGGAGGCTACTATGT  
GTATCGTCTGACCAAGCCCAGCGTCTGCTTCCACGTCTACTGTGGTCATTTTTATGACATCTGCG  
ACGAGGACTGCCATGGCAGCTGCTCAGATACCAGCGAGTGCACATGCGCTCCAGGAACTGTGCTA  
GGCCCTGACAGGCAGACATGCTTTGATGAAAATGAATGTGAGCAAAACAACGGTGGCTGCAGTGA  
GATCTGTGTGAACCTCAAAAACCTCCTACCGCTGTGAGTGTGGGGTTGGCCGTGTGCTAAGAAGTG  
ATGGCAAGACTTGTGAAGACGTTGAAGGATGCCACAATAACAATGGTGGCTGCAGCCACTCTTGC  
CTTGATCTGAGAAAGGCTACCAAGTGTGAATGTCCCCGGGGCCTGGTGCTGTCTGAGGATAACCA  
CACTTGCCAAGTCCCTGTGTTGTGCAAATCAAATGCCATTGAAGTGAACATCCCCAGGGAGCTGG  
TTGGTGGCCTGGAGCTCTTCCTGACCAACACCTCCTGCCGAGGAGTGTCCAACGGCACCCATGTC  
AACATCCTCTTCTCTCTCAAGACATGTGGTACAGTGGTTCGATGTGGTGAATGACAAGATTGTGGC  
CAGCAACCTCGTGACAGGTCTACCCAAGCAGACCCCGGGGAGCAGCGGGGACTTCATCATCCGAA  
CCAGCAAGCTGCTGATCCCGGTGACCTGCGAGTTTCCACGCCTGTACACCATTTCTGAAGGATAC  
GTTCCCAACCTTCGAAACTCCCCACTGGAAATCATGAGCCGAAATCATGGGATCTTCCCATTAC  
TCTGGAGATCTTCAAGGACAATGAGTTTGAAGAGCCTTACCGGGAAGCTCTGCCCACCCTCAAGC  
TTCGTGACTCCCTCTACTTTGGCATTGAGCCCGTGGTGCACGTGAGCGGCTTGGAAAGCTTGGTG  
GAGAGCTGCTTTGCCACCCCCACCTCCAAGATCGACGAGGTCTGAAATACTACCTCATCCGGGA  
TGGCTGTGTTTCAGATGACTCGGTAAAGCAGTACACATCCCGGGATCACCTAGCAAAGCACTTCC  
AGGTCCTGTCTTCAAGTTTGTGGGCAAAGACCACAAGGAAGTGTCTGCACTGCCGGGTCTT  
GTCTGTGGAGTGTTGGACGAGCGTTCCCGCTGTGCCAGGGTTGCCACCGGCGAATGCGTCGTGG  
GGCAGGAGGAGAGGACTCAGCCGGTCTACAGGGCCAGACGCTAACAGGCGGCCCGATCCGCATCG  
ACTGGGAGGACT**TAG**TTTCGTAGCCATACCTCGAGTCCCTGCATTGGACGGCTCTGCTCTTTGGAGC  
TTCTCCCCCACC GCCCTCTAAGAACATCTGCCAACAGCTGGGTTTACAGACTTCACACTGTGAGTT  
CAGACTCCCAGCACCAACTCACTCTGATTCTGGTCCATTACAGTGGGCACAGGTCACAGCACTGCT  
GAACAATGTGGCCTGGGTGGGGTTTCATCTTTCTAGGGTTGAAAACCTAACTGTCCACCCAGAAA  
GACACTCACCCCATTTCCCTCATTTCTTTCCCTACACTTAAATACCTCGTGTATGGTGCAATCAGA  
CCACAAAATCAGAAGCTGGGTATAATATTTCAAGTTACAAACCCTAGAAAAATTAACAGTTACT  
GAAATTATGACTTAAATACCCAATGACTCCTTAAATATGTAAATTATAGTTATACCTTGAAATTT  
CAATTCAAATGCAGACTAATTATAGGGAATTTGGAAGTGTATCAATAAAACAGTATATAATTTT

## **FIGURE 110**

MPPFLLLTCLFITGTSVSPVALDPCSAYISLNEPWRNTDHQLDESQGPPLCDNHVNGEWYHFTGM  
AGDAMPTFCIPENHCGTHAPVWLNGSHPLEGDGIVQRQACASFNGNCCLWNTTVEVKACPGGYV  
YRLTKPSVCFHVYCGHFYDICDEDCHGSCSDTSECTCAPGTVLGPDRQTCFDENECEQNNGGCSE  
ICVNLKNSYRCECGVGRVLRSDGKTCEDEVEGCHNNNGGCSHSCLGSEKGYQCECPRGLVLSEDNH  
TCQVPVLCKSNAIEVNIPRELVGGLLEFLTNTSCRGVSNNGTHVNILFSLKTCGTVVDVVNDKIVA  
SNLVTGLPKQTPGSSGDFIIRTSKLLIPVTCEFPRLYTISEGYVPNLRNSPLEIMSRNHGIFPFT  
LEIFKDNEFEOPYREALPTLKLRLDSLYFGIEPVVHVSGLESLEVSCFATPTSKIDEVLKYYLIRD  
GCVSDDSVKQYTSRDHLAKHFQVPVFKFVGKDHKEVFLHCRVLVCGVLDERSRCAQGCHRRMRRG  
AGGEDSAGLQGQTLTGGPIDWED

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-16

#### **N-glycosylation sites.**

amino acids 89-93, 116-120, 259-263, 291-295, 299-303

#### **Tyrosine kinase phosphorylation sites.**

amino acids 411-418, 443-451

#### **N-myristoylation sites.**

amino acids 226-232, 233-239, 240-246, 252-258, 296-302, 300-306,  
522-528, 531-537

#### **Aspartic acid and asparagine hydroxylation site.**

amino acids 197-209

#### **ZP domain proteins.**

amino acids 431-457

#### **Calcium-binding EGF-like proteins.**

amino acids 191-212, 232-253



**FIGURE 111**

GAGAGAGGCAGCAGCTTGCTCAGCGGACAAGGATGCTGGGCGTGAGGGACCAAGGCCTGCCCTGC  
ACTCGGGCCTCCTCCAGCCAGTGCTGACCAGGGACTTCTGACCTGCTGGCCAGCCAGGACCTGTG  
TGGGGAGGCCCTCCTGCTGCCCTTGGGGTGACAATCTCAGCTCCAGGCTACAGGGAGACCGGGAGG  
ATCACAGAGCCAGC**ATG**TTACAGGATCCTGACAGTGATCAACCTCTGAACAGCCTCGATGTCAAA  
CCCCTGCGCAAACCCCGTATCCCCATGGAGACCTTCAGAAAGGTGGGGATCCCCATCATCATAGC  
ACTACTGAGCCTGGCGAGTATCATCATTGTGGTTGTCCTCATCAAGGTGATTCTGGATAAATACT  
ACTTCCTCTGCGGGCAGCCTCTCCACTTCATCCCGAGGAAGCAGCTGTGTGACGGAGAGCTGGAC  
TGTCCCTTGGGGGAGGACGAGGAGCACTGTGTCAAGAGCTTCCCCGAAGGGCCTGCAGTGGCAGT  
CCGCCTCTCCAAGGACCGATCCACACTGCAGGTGCTGGACTCGGCCACAGGGAACTGGTTCTCTG  
CCTGTTTCGACAACTTCACAGAAGCTCTCGCTGAGACAGCCTGTAGGCAGATGGGCTACAGCAGA  
GCTGTGGAGATTGGCCAGACCAGGATCTGGATGTTGTTGAAATCACAGAAAACAGCCAGGAGCT  
TCGCATGCGGAACTCAAGTGGGCCCTGTCTCTCAGGCTCCCTGGTCTCCCTGCACTGTCTTGCCT  
GTGGGAAGAGCCTGAAGACCCCCGTGTGGTGGGTGGGGAGGAGGCCTCTGTGGATTCTTGGCCT  
TGGCAGGTCAGCATCCAGTACGACAAACAGCACGTCTGTGGAGGGAGCATCCTGGACCCCCACTG  
GGTCCTCACGGCAGCCCACTGCTTCAGGAAACATAACCGATGTGTTCAACTGGAAGGTGCGGGCAG  
GCTCAGACAAACTGGGCAGCTTCCCATCCCTGGCTGTGGCCAAGATCATCATCATTGAATTCAAC  
CCCATGTACCCCAAAGACAATGACATCGCCCTCATGAAGCTGCAGTTCCCACTCACTTTCTCAGG  
CACAGTCAGGCCCATCTGTCTGCCCTTCTTTGATGAGGAGCTCACTCCAGCCACCCCACTCTGGA  
TCATTGGATGGGGCTTTACGAAGCAGAATGGAGGGAAGATGTCTGACATACTGCTGCAGGCGTCA  
GTCCAGGTCATTGACAGCACACGGTGCAATGCAGACGATGCGTACCAGGGGGAAGTCACCGAGAA  
GATGATGTGTGCAGGCATCCCGGAAGGGGGTGTGGACACCTGCCAGGGTGACAGTGGTGGGCCCC  
TGATGTACCAATCTGACCAGTGGCATGTGGTGGGCATCGTTAGCTGGGGCTATGGCTGCGGGGGC  
CCGAGCACCCCAGGAGTATACACCAAGGTCTCAGCCTATCTCAACTGGATCTACAATGTCTGGAA  
GGCTGAGCTG**TAA**TGCTGCTGCCCCCTTTCAGTGCTGGGAGCCGCTTCCTTCCTGCCCTGCCAC  
CTGGGGATCCCCCAAAGTCAGACACAGAGCAAGAGTCCCCTTGGGTACACCCCTCTGCCACAGC  
CTCAGCATTTCTTGGAGCAGCAAAGGGCCTCAATTCCCTGTAAGAGACCCTCGCAGCCCAGAGGCG  
CCCAGAGGAAGTCAGCAGCCCTAGCTCGGCCACACTTGGTGCTCCCAGCATCCCAGGGAGAGACA  
CAGCCCACTGAACAAGGTCTCAGGGGTATTGCTAAGCCAAGAAGGAACTTTCCACACTACTGAA  
TGGAAGCAGGCTGTCTTGTAAGAGCCAGATCACTGTGGGCTGGAGAGGAGAAGGAAAGGGTCTG  
CGCCAGCCCTGTCCGTCTTCACCCATCCCCAAGCCTACTAGAGCAAGAAACCAGTTGTAATATAA  
AATGCACTGCCCTACTGTTGGTATGACTACCGTTACCTACTGTTGTCATTGTTATTACAGCTATG  
GCCACTATTATTAAAGAGCTGTGTAACATCTCTGGCAAAAAAAAAAAAA

112/168

## **FIGURE 112**

MLQDPDSQPLNSLDVKPLRKPRIPMETFRKVGIPIIIALLSLASIIIVVVLIKVILDKYYFLCG  
QPLHFIPRKQLCDGELDCPLGEDEEHCVKSFPEGPAVAVRLSKDRSTLQVLDSATGNWFSACFDN  
FTEALAETACRQMGYSRAVEIGPDQDLDVVEITENSQELMRNSSGPCLSGSLVSLHCLACGKSL  
KTPRVVGEEASVDSWPWQVSIQYDKQHVCGGSILDPHWLTAAHCFRKHTDVFNWKVRAGSDKL  
GSFPSLAVAKIIIIIEFNPMYPKDNDIALMKLQFPLTFSGTVRPICLPFFDEELTPATPLWIIGWG  
FTKQNGGKMSDILLQASVQVIDSTRCNADDAYQGEVTEKMMCAGIPEGGVDTCCQGDGGPLMYQS  
DQWHVVGIVSWGYGCGGPSTPGVYTKVSAYLNWIYNVWKAEL

### **Transmembrane domain:**

amino acids 32-53 (typeII)

**FIGURE 113**

GGCTGGACTGGAACCTCCTGGTCCCAAGTGATCCACCCGCCTCAGCCTCCCAAGGTGCTGTGATTA  
TAGGTGTAAGCCACCGTGTCTGGCCTCTGAACAACCTTTTTCAGCAACTAAAAAGCCACAGGAGT  
TGAACCTGCTAGGATTCTGACTATGCTGTGGTGGCTAGTGCTCCTACTCCTACCTACATTAAAATC  
TGTTTTTTGTTCTCTTGTAAC TAGCCTTTACCTTCCTAACACAGAGGATCTGTCACTGTGGCTCT  
GGCCCAAACCTGACCTTCACTCTGGAACGAGAACAGAGGTTTCTACCCACACCGTCCCCTCGAAG  
CCGGGGACAGCCTCACCTTGCTGGCCTCTCGCTGGAGCAGTGCCCTACCAACTGTCTCACGTCT  
GGAGGCACTGACTCGGGCAGTGCAGGTAGCTGAGCCTCTTGGTAGCTGCGGCTTTCAAGGTGGGC  
CTTGCCCTGGCCGTAGAAGGGATTGACAAGCCCGAAGATTTTCATAGGCGATGGCTCCCCTGCCC  
AGGCATCAGCCTTGCTGTAGTCAATCACTGCCCTGGGGCCAGGACGGGCGTGGACACCTGCTCA  
GAAGCAGTGGGTGAGACATCACGCTGCCCGCCCATCTAACCTTTTCATGTCTGCACATCACCTG  
ATCCATGGGCTAATCTGAACTCTGTCCCAAGGAACCCAGAGCTTGAGTGAGCTGTGGCTCAGACC  
CAGAAGGGGTCTGCTTAGACCACCTGGTTTATGTGACAGGACTTGCATTCTCCTGGAACATGAGG  
GAACGCCGGAGGAAAGCAAAGTGGCAGGGAAGGAACTTGTGCCAAATTATGGGTGAGAAAAGATG  
GAGGTGTTGGGTATCACAAGGCATCGAGTCTCCTGCATTCACTGGACATGTGGGGGAAGGGCTG  
CCGATGGCGCATGACACACTCGGGACTCACCTCTGGGGCCATCAGACAGCCGTTTCCGCCCCGAT  
CCACGTACCAGCTGCTGAAGGGCAACTGCAGGCCGATGCTCTCATCAGCCAGGCAGCAGCCAAAA  
TCTGCGATCACCAGCCAGGGGCAGCCGTCTGGGAAGGAGCAAGCAAAGTGACCATTTCTCCTCCC  
CTCCTTCCCTCTGAGAGGCCCTCCTATGTCCCTACTAAAGCCACCAGCAAGACATAGCTGACAGG  
GGCTAATGGCTCAGTGTTGGCCCAGGAGGTCAGCAAGGCCTGAGAGCTGATCAGAAGGGCCTGCT  
GTGCGAACACGGAAATGCCTCCAGTAAGCACAGGCTGCAAAATCCCCAGGCAAAGGACTGTGTGG  
CTCAATTTAAATCATGTTCTAGTAATTGGAGCTGTCCCCAAGACCAAAGGAGCTAGAGCTTGGTT  
CAAATGATCTCCAAGGGCCCTTATACCCCAGGAGACTTTGATTTGAATTTGAAACCCCAAATCCA  
AACCTAAGAACCAGGTGCATTAGAATCAGTTATTGCCGGGTGTGGTGGCCTGTAATGCCAACAT  
TTTGGGAGGCCGAGGCGGGTAGATCACCTGAGGTCAGGAGTTCAAGACCAGCCTGGCCAACATGG  
TGAAACCCCTGTCTCTACTAAAAATACAAAAAACTAGCCAGGCATGGTGGTGTGTGCCTGTATC  
CCAGCTACTCGGGAGGCTGAGACAGGAGAATTACTTGAACCTGGGAGGTGAAGGAGGCTGAGACA  
GGAGAATCACTTCAGCCTGAGCAACACAGCGAGACTCTGTCTCAGAAAAATAAAAAAGAATTA  
TGGTTATTTGTAA

114/168

## **FIGURE 114**

MLWWLVLLLLPTLKSVFCSLVTSLYLPNTEDLSLWLWPKPDLHSGTRTEVSTHTVPSKPGTASPC  
WPLAGAVPSPTVSRLEALTRAVQVAEPLGSCGFQGGPCPGRRRD

**Signal peptide:**

amino acids 1-15

**FIGURE 115**

CAGCAGTGGTCTCTCAGTCCTCTCAAAGCAAGGAAAGAGTACTGTGTGCTGAGAGACC**ATGGCAA**  
AGAATCCTCCAGAGAATTGTGAAGACTGTCACATTCTAAATGCAGAAGCTTTTAAATCCAAGAAA  
ATATGTAAATCACTTAAGATTTGTGGACTGGTGTGTTGGTATCCTGGCCCTAACTCTAATTGTCCT  
GTTTTGGGGGAGCAAGCACTTCTGGCCGGAGGTACCCAAAAAGCCTATGACATGGAGCACACTT  
TCTACAGCAATGGAGAGAAGAAGAAGATTTACATGGAAATTGATCCTGTGACCAGAACTGAAATA  
TTCAGAAGCGGAAATGGCACTGATGAAACATTGGAAGTGCACGACTTTAAAAACGGATACACTGG  
CATCTACTTCGTGGGTCTTCAAAAATGTTTTATCAAACTCAGATTAAAGTGATTTCCTGAATTTT  
CTGAACCAGAAGAGGAAATAGATGAGAATGAAGAAATTACCACAACCTTTCTTTGAACAGTCAGTG  
ATTTGGGTCCCAGCAGAAAAGCCTATTGAAAACCGAGATTTTCTTAAAAATTCCAAAATTCTGGA  
GATTTGTGATAACGTGACCATGTATTGGATCAATCCCACTCTAATATCAGTTTCTGAGTTACAAG  
ACTTTGAGGAGGAGGGAGAAGATCTTCACTTTCCTGCCAACGAAAAAAAGGGATTGAACAAAAT  
GAACAGTGGGTGGTCCCTCAAGTGAAAGTAGAGAAGACCCGTCACGCCAGACAAGCAAGTGAGGA  
AGAACTTCCAATAAATGACTATACTGAAAATGGAATAGAATTTGATCCCATGCTGGATGAGAGAG  
GTTATTGTTGTATTTACTGCCGTCGAGGCAACCGCTATTGCCGCCGCGTCTGTGAACCTTTACTA  
GGCTACTACCCATATCCATACTGCTACCAAGGAGGACGAGTCATCTGTCGTGTCATCATGCCTTG  
TAACTGGTGGGTGGCCCGCATGCTGGGGAGGGTCT**TAA**TAGGAGGTTTGAGCTCAAATGCTTAAAC  
TGCTGGCAACATATAATAAATGCATGCTATTCAATGAATTTCTGCCTATGAGGCATCTGGCCCT  
GGTAGCCAGCTCTCCAGAATTACTTGTAGGTAATTCCTCTCTTCATGTTCTAATAAACTTCTACA  
TTATCACCAAAAAAAAAAAAAAAAAAAAAA

## **FIGURE 116**

MAKNPPENCEDCHILNAEAFKSKKICKSLKICGLVFGILALT LIVLFWGSKHFWPEVPKKAYDME  
HTFYSNGEKKKIYMEIDPVTRTEIFRSGNGTDETLEVHDFKNGYTG IYFVGLQKCFIKTQIKVIP  
EFSEPEEEIDENEEITTTFFEQSVIWVPAEKPIENRDFLKN SKILEICDNVTMYWINPTLISVSE  
LQDFEEEGEDLHF PANEEKGIEQNEQWVVPQVKVEKTRHARQASEEELPINDYTENGIEFDPMLD  
ERGYCCIIYCRRGNRYCRRVCEPLLGYYPYPYCYQGGRVICRVIMPCNWWVARMLGRV

**Important features of the protein:**

**Signal peptide:**

amino acids 1-40

**Transmembrane domain:**

amino acids 25-47 (type II)

**N-glycosylation sites.**

amino acids 94-97, 180-183

**Glycosaminoglycan attachment sites.**

amino acids 92-95, 70-73, 85-88, 133-136, 148-151, 192-195, 239-242

**N-myristoylation sites.**

amino acids 33-38, 95-100, 116-121, 215-220, 272-277

**Microbodies C-terminal targeting signal.**

amino acids 315-317

**Cytochrome c family heme-binding site signature.**

amino acids 9-14

**FIGURE 117**

GAGCTCCCCTCAGGAGCGCGTTAGCTTCACACCTTCGGCAGCAGGAGGGCGGCAGCTTCTCGCAG  
GCGGCAGGGCGGGCGGCCAGGATC**ATGT**TCCACCACCACATGCCAAGTGGTGGCGTTCCCTCCTGTC  
CATCCTGGGGCTGGCCGGCTGCATCGCGGCCACCGGGATGGACATGTGGAGCACCCAGGACCTGT  
ACGACAACCCCGTCACCTCCGTGTTCCAGTACGAAGGGCTCTGGAGGAGCTGCGTGAGGCAGAGT  
TCAGGCTTCACCGAATGCAGGCCCTATTTACCATCCTGGGACTTCCAGCCATGCTGCAGGCAGT  
GCGAGCCCCTGATGATCGTAGGCATCGTCCTGGGTGCCATTGGCCTCCTGGTATCCATCTTTGCCC  
TGAAATGCATCCGCATTGGCAGCATGGAGGACTCTGCCAAAGCCAACATGACACTGACCTCCGGG  
ATCATGTTTATTGTCTCAGGTCTTTGTGCAATTGCTGGAGTGTCTGTGTTTGCCAACATGCTGGT  
GACTAACTTCTGGATGTCCACAGCTAACATGTACACCGGCATGGGTGGGATGGTGCAGACTGTTT  
AGACCAGGTACACATTTGGTGCGGCTCTGTTTCGTGGGCTGGGTGCGTGGAGGCCTCACACTAATT  
GGGGGTGTGATGATGTGCATCGCCTGCCGGGGCCTGGCACCAGAAGAAACCAACTACAAAGCCGT  
TTCTTATCATGCCTCAGGCCACAGTGTTGCCTACAAGCCTGGAGGCTTCAAGGCCAGCACTGGCT  
TTGGGTCCAACACCAAAAACAAGAAGATATACGATGGAGGTGCCCCGACAGAGGACGAGGTACAA  
TCTTATCCTTCCAAGCACGACTATGTG**TAA**TGCTCTAAGACCTCTCAGCACGGGCGGAAGAACT  
CCCGGAGAGCTCACCCAAAAACAAGGAGATCCCATCTAGATTTCTTCTTGCTTTTGACTCACAG  
CTGGAAGTTAGAAAAGCCTCGATTTTCATCTTTGGAGAGGCCAAATGGTCTTAGCCTCAGTCTCTG  
TCTCTAAATATTCCACCATAAAACAGCTGAGTTATTTATGAATTAGAGGCTATAGCTCACATTTT  
CAATCCTCTATTTCTTTTTTTTAAATATAACTTTCTACTCTGATGAGAGAATGTGGTTTTTAATCTC  
TCTCTCACATTTTGATGATTTAGACAGACTCCCCCTCTCCTCCTAGTCAATAAACCCATTGATG  
ATCTATTTCCCAGCTTATCCCCAAGAAAACCTTTTGAAAGGAAAGAGTAGACCCAAAGATGTTATT  
TTCTGCTGTTTGAATTTTGTCTCCCCACCCCCAAGTTGGCTAGTAATAAACACTTACTGAAGAAG  
AAGCAATAAGAGAAAGATATTTGTAATCTCTCCAGCCCATGATCTCGGTTTTCTTACACTGTGAT  
CTTAAAAGTTACCAAACCAAAGTCATTTTCAGTTTGAGGCAACCAAACCTTTCTACTGCTGTTGA  
CATCTTCTTATTACAGCAACACCATTCTAGGAGTTTCCTGAGCTCTCCACTGGAGTCCCTCTTCT  
GTCGCGGGTCAGAAATTGTCCCTAGATGAATGAGAAAATTATTTTTTTTAAATTTAAGTCCTAAAT  
ATAGTTAAAATAAATAATGTTTTTAGTAAAATGATACACTATCTCTGTGAAATAGCCTCACCCCTA  
CATGTGGATAGAAGGAAATGAAAAATAATTGCTTTGACATTGTCTATATGGTACTTTGTAAAGT  
CATGCTTAAGTACAAATTCCATGAAAAGCTCACACCTGTAATCCTAGCACTTTGGGAGGCTGAGG  
AGGAAGGATCACTTGAGCCCAGAAGTTCGAGACTAGCCTGGGCAACATGGAGAAGCCCTGTCTCT  
ACAAAATACAGAGAGAAAAATCAGCCAGTCATGGTGGCATAACCTGTAGTCCCAGCATTTCCGG  
GAGGCTGAGGTGGGAGGATCACTTGAGCCCAGGGAGGTTGGGGCTGCAGTGAGCCATGATCACAC  
CACTGCACTCCAGCCAGGTGACATAGCGAGATCCTGTCTAAAAAAATAAAAAATAAATAATGGAA  
CACAGCAAGTCCTAGGAAGTAGGTTAAAACATAATTCTTTAA

## **FIGURE 118**

MSTTTCQVVAFLLSILGLAGCIAATGMDMWSTQDLYDNPVTSVFQYEGLRSCVRQSSGFTECRP  
YFTILGLPAMLQAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLTSGIMFIVSGL  
CAIAGVSVFANMLVTNFWMSTANMYTGMGGMVQTVQTRYTFGAALFVGWVAGGLTLIGGVMMCIA  
CRGLAPEETNYKAVSYHASGHSVAYKPGGFKASTGFGSNTKNKKIYDGGARTEDEVQSYPSKHDY  
V

**Signal peptide:**

amino acids 1-23

**Transmembrane domains:**

amino acids 81-100, 121-141, 173-194



**FIGURE 119**

GGAAAACTGTTCTCTTCTGTGGCACAGAGAACCCTGCTTCAAAGCAGAAGTAGCAGTTCCGGAG  
 TCCAGCTGGCTAAAACATCCCAGAGGATA**ATG**GCAACCCATGCCTTAGAAATCGCTGGGCTGT  
 TTCTTGGTGGTGTGGAATGGTGGGCACAGTGGCTGTCACTGTCAATGCCTCAGTGGAGAGTGTCTG  
 GCCTTCATTGAAAACAACATCGTGGTTTTTTGAAAACCTTCTGGGAAGGACTGTGGATGAATTGCGT  
 GAGGCAGGCTAACATCAGGATGCAGTGCAAAATCTATGATTCCCTGCTGGCTCTTTCTCCGGACC  
 TACAGGCAGCCAGAGGACTGATGTGTGCTGCTTCCGTGATGTCCTTCTTGGCTTTCATGATGGCC  
 ATCCTTGGCATGAAATGCACCAGGTGCACGGGGGACAATGAGAAGGTGAAGGCTCACATTCTGCT  
 GACGGCTGGAATCATCTTCATCATCACGGGCATGGTGGTGTCTATCCCTGTGAGCTGGGTGGCA  
 ATGCCATCATCAGAGATTTCTATAACTCAATAGTGAATGTTGCCCAAAAACGTGAGCTTGGAGAA  
 GCTCTCTACTTAGGATGGACCACGGCACTGGTGGTGTGATGTTGGAGGAGCTCTGTTCTGCTGCGT  
 TTTTTGTTGCAACGAAAAGAGCAGTAGCTACAGATACTCGATACCTTCCCATCGCACAAACCCAAA  
 AAAGTTATCACACCGGAAAGAAGTCACCGAGCGTCTACTCCAGAAGTCAGTATGT**TAG**TTGTGT  
 ATGTTTTTTTAACTTTACTATAAAGCCATGCAAATGACAAAAATCTATATTACTTTCTCAAAATG  
 GACCCCAAAGAACTTTGATTTACTGTTCTTAACTGCCTAATCTTAATTACAGGAAGTGTGCATC  
 AGCTATTTATGATTCTATAAGCTATTTTCAGCAGAATGAGATATTAAACCAATGCTTTGATTGTT  
 CTAGAAAGTATAGTAATTTGTTTTCTAAGGTGGTTCAAGCATCTACTCTTTTATCATTTACTTC  
 AAAATGACATTGCTAAAGACTGCATTATTTTACTACTGTAATTTCTCCACGACATAGCATTATGT  
 ACATAGATGAGTGTAACATTTTATATCTCACATAGAGACATGCTTATATGGTTTTATTTAAAATGA  
 AATGCCAGTCCATTACACTGAATAAATAGAACTCAACTATTGCTTTTCAGGGAAATCATGGATAG  
 GGTGGAAGAAGGTTACTATTAATTGTTTTAAAACAGCTTAGGGATTAATGTCTCCATTTTATAAT  
 GAAGATTAAATGAAGGCTTTAATCAGCATTGTAAAGGAAATTGAATGGCTTTCTGATATGCTGT  
 TTTTAGCCTAGGAGTTAGAAATCCTAACTTCTTTATCCTCTTCTCCCAGAGGCTTTTTTTTTCT  
 TGTGTATTAAATTAACATTTTTTAAAACGCAGATATTTTGTCAAGGGGCTTTGCATTCAAACCTGCT  
 TTTCCAGGGCTATACTCAGAAGAAAGATAAAAGTGTGATCTAAGAAAAAGTGATGGTTTTAGGAA  
 AGTGAAAATATTTTTGTTTTTGTATTTGAAGAAGAATGATGCATTTTGACAAGAAATCATATATG  
 TATGGATATATTTTAATAAGTATTTGAGTACAGACTTTGAGGTTTCATCAATATAAATAAAAGAG  
 CAGAAAAATATGTCTTGGTTTTTCATTTGCTTACCAAAAAACAACAACAAAAAAGTTGTCCTTT  
 GAGAACTTCACCTGCTCCTATGTGGGTACCTGAGTCAAAATTGTCATTTTTGTTCTGTGAAAAAT  
 AAATTTCCCTTCTTGTACCATTTCTGTTTAGTTTTACTAAAATCTGTAAATACTGTATTTTTCTGT  
 TTATTCCAAATTTGATGAACTGACAATCCAATTTGAAAGTTTGTGTGACGTCTGTCTAGCTTA  
 AATGAATGTGTTCTATTTGCTTTATACATTTATATTAATAAATTGTACATTTTTCTAATT

120/168

## **FIGURE 120**

MATHALEIAGLFLGGVGMVGTVAVTVMPPQWRVSAFIENNIVVFENFW EGLWMNCVRQANIRMQCK  
IYDSLLALSPDLQAARGLMCAASVMSFLAFMMAILGMKCTRCTGDNEKVKAHILLTAGIIFIITG  
MVLIPVSWVANAIIRDFYNSIVNVAQKRELGEALYLGWTTALVLIVGGALFCCVFCCNEKSSSY  
RYSIPSHRTTQKSYHTGKKSPSVYSRSQYV

### **Signal peptide:**

amino acids 1-17

### **Transmembrane domains:**

amino acids 82-101, 118-145, 164-188

**FIGURE 121**

GGAGAGAGGCGCGCGGGTGAAAGGCGCATTGATGCAGCCTGCGGCGGCCTCGGAGCGCGGCGGAG  
CCAGACGCTGACCACGTTCTCTCCTCGGTCTCCTCCGCCTCCAGCTCCGCGCTGCCCCGGCAGCC  
GGGAGCCATGCGACCCCAGGGCCCCGCGCCTCCCCGCAGCGGCTCCGCGGCCTCCTGCTGCTCC  
TGCTGCTGCAGCTGCCCCGCGCCGTGAGCGCCTCTGAGATCCCCAAGGGGAAGCAAAAGGCGCAG  
CTCCGGCAGAGGGAGGTGGTGGACCTGTATAATGGAATGTGCTTACAAGGGCCAGCAGGAGTGCC  
TGGTCGAGACGGGAGCCCTGGGGCCAATGTTATTCCGGGTACACCTGGGATCCCAGGTCGGGATG  
GATTCAAAGGAGAAAAGGGGGAATGTCTGAGGGAAAGCTTTGAGGAGTCCTGGACACCCAACTAC  
AAGCAGTGTTTCATGGAGTTCATTGAATTATGGCATAGATCTTGGGAAAATTGCGGAGTGACATT  
TACAAAGATGCGTTCAAATAGTGCTCTAAGAGTTTTGTTTCAGTGGCTCACTTCGGCTAAAATGCA  
GAAATGCATGCTGTCTAGCGTTGGTATTTACATTCAATGGAGCTGAATGTTTCAGGACCTCTTCCC  
ATTGAAGCTATAATTTATTTGGACCAAGGAAGCCCTGAAATGAATTCAACAATTAATATTCATCG  
CACTTCTTCTGTGGAAGGACTTTGTGAAGGAATTGGTGCTGGATTAGTGGATGTTGCTATCTGGG  
TTGGCACTTGTTTCAGATTACCCAAAAGGAGATGCTTCTACTGGATGGAATTCAGTTTCTCGCATC  
ATTATTGAAGAACTACCAAAATAAATGCTTTAATTTTCATTTGCTACCTCTTTTTTTATTATGCC  
TTGGAATGGTTCACTTAAATGACATTTTAAATAAGTTTATGTATACATCTGAATGAAAAGCAAAG  
CTAAATATGTTTACAGACCAAAGTGTGATTTCACACTGTTTTTAAATCTAGCATTATTCATTTTG  
CTTCAATCAAAGTGGTTTCAATATTTTTTTTAGTTGGTTAGAATACTTTCTTCATAGTCACATT  
CTCTCAACCTATAATTTGGAATATTGTTGTGGTCTTTTGTTTTTTCTCTTAGTATAGCATTTTTTA  
AAAAATATAAAAGCTACCAATCTTTGTACAATTTGTAAATGTAAAGAATTTTTTTTTATATCTGT  
TAAATAAAAATTATTTCCAACA

122/168

## **FIGURE 122**

MRPQGPAASPQRLRGLLLLLLLQLPAPSSASEIPKGKQKAQLRQREVVDLYNGMCLQGPAGVPGR  
DGSPGANVIPGTPGIPGRDGFKEKGECLRESFEESWTPNYKQCSWSSLNYGIDLGKIAECTFTK  
MRSNSALRVLFSGSLRLKCRNACCQRWYFTFNGAECSGPLPIEAIYLDQGSPENNSTINIHRTS  
SVEGLCEGIGAGLVDVAIWVGTCSDYPKGDASTGWNSVSRIIEELPK

### **Signal peptide:**

amino acids 1-30

### **Transmembrane domain:**

amino acids 195-217

**FIGURE 123**

GCTGAGCGTGTGCGCGGTACGGGGCTCTCCTGCCTTCTGGGCTCCAACGCAGCTCTGTGGCTGAA  
CTGGGTGCTCATCACGGGAAGTCTGGGCTATGGAATACAGATGTGGCAGCTCAGGTAGCCCCAA  
ATTGCCTGGAAGAATACATCATGTTTTTCGATAAGAAGAAATTGTAGGATCCAGTTTTTTTTTTA  
ACCGCCCCCTCCCCACCCCCCAAAAAAAGTGTAAAGATGCAAAAACGTAATATCCATGAAGATCC  
TATTACCTAGGAAGATTTTGATGTTTTGCTGCGAATGCGGTGTTGGGATTTATTTGTTCTTGGAG  
TGTTCTGCGTGGCTGGCAAAGAATAATGTTCCAAAATCGGTCCATCTCCCAAGGGGTCCAATTTT  
TCTTCTGGGTGTCAGCGAGCCCTGACTCACTACAGTGCAGCTGACAGGGGGTGTGATGCAACTG  
GCCCCTAAGCCAAAGCAAAAGACCTAAGGACGACCTTTGAACAATACAAAGG**ATG**GGTTTCAATG  
TAATTAGGCTACTGAGCGGATCAGCTGTAGCACTGGTTATAGCCCCACTGTCTTACTGACAATG  
CTTTCTTCTGCCGAACGAGGATGCCCTAAGGGCTGTAGGTGTGAAGGCAAAATGGTATATTGTGA  
ATCTCAGAAATTACAGGAGATACCCTCAAGTATATCTGCTGGTTGCTTAGGTTTGTCCCTTCGCT  
ATAACAGCCTTCAAAAAGTAAAGTATAATCAATTTAAAGGGCTCAACCAGCTCACCTGGCTATAC  
CTTGACCATAACCATATCAGCAATATTGACGAAAATGCTTTTAAATGGAATACGCAGACTCAAAGA  
GCTGATTCTTAGTTCCAATAGAATCTCCTATTTTCTTAACAATACCTTCAGACCTGTGACAAATT  
TACGGAAGCTTGGATCTGTCTTATAATCAGCTGCATTCTCTGGGATCTGAACAGTTTCGGGGCTTG  
CGGAAGCTGCTGAGTTTACATTTACGGTCTAACTCCCTGAGAACCATCCCTGTGCGAATATTCCA  
AGACTGCCGCAACCTGGAAGTTTTGGACCTGGGATATAACCGGATCCGAAGTTTAGCCAGGAATG  
TCTTTGCTGGCATGATCAGACTCAAAGAAGTTTACCTGGAGCACAATCAATTTTCCAAGCTCAAC  
CTGGCCCTTTTTCCAAGGTTGGTCAGCCTTCAGAACCTTTACTTGCAGTGGAATAAAATCAGTGT  
CATAGGACAGACCATGTCTTGGACCTGGAGCTCCTTACAAAGGCTTGATTTATCAGGCAATGAGA  
TCGAAGCTTTCAGTGGACCCAGTGTTTTCCAGTGTGTCCCGAATCTGCAGCGCCTCAACCTGGAT  
TCCAACAAGCTCACATTTATTGGTCAAGAGATTTTGGATTCTTGGATATCCCTCAATGACATCAG  
TCTTGCTGGGAATATATGGGAATGCAGCAGAAATATTTGCTCCCTTGTAAGTGGCTGAAAAGTT  
TTAAAGGTCTAAGGGAGAATACAATTATCTGTGCCAGTCCCAAAGAGCTGCAAGGAGTAAATGTG  
ATCGATGCAGTGAAGAACTACAGCATCTGTGGCAAAAGTACTACAGAGAGGTTTGATCTGGCCAG  
GGCTCTCCCAAAGCCGACGTTTAAGCCCAAGCTCCCCAGGCCGAAGCATGAGAGCAAACCCCTT  
TGCCCCCGACGGTGGGAGCCACAGAGCCCGGCCAGAGACCGATGCTGACGCCGAGCACATCTCT  
TTCCATAAAATCATCGCGGGCAGCGTGGCGCTTTTCTGTCCGTGCTCGTCATCCTGCTGGTTAT  
CTACGTGTCATGGAAGCGGTACCCTGCGAGCATGAAGCAGCTGCAGCAGCGCTCCCTCATGCGAA  
GGCACAGGAAAAAGAAAAGACAGTCCCTAAAGCAAATGACTCCCAGCACCCAGGAATTTTATGTA  
GATTATAAACCCACCAACACGGAGACCAGCGAGATGCTGCTGAATGGGACGGGACCCTGCACCTA  
TAACAAATCGGGCTCCAGGGAGTGTGAGGTAT**TGA**ACCATTGTGATAAAAAGAGCTCTTAAAAGCT  
GGGAAATAAGTGGTGCTTTATTGAACTCTGGTGACTATCAAGGGAACGCGATGCCCCCTCCCC  
TTCCCTCTCCCTCTCACTTTGGTGGCAAGATCCTTCCCTTGTCCGTTTTAGTGCATTCATAATACT  
GGTCATTTTCTCTCATACATAATCAACCCATTGAAATTTAAATACCACAATCAATGTGAAGCTT  
GAACTCCGGTTTAATATAATACCTATTGTATAAGACCCTTTACTGATTCCATTAATGTCGCATTT  
GTTTTAAGATAAAAAGTCTTTTCATAGGTAAAAA

## **FIGURE 124**

MGFNVIRLLSGSAVALVIAPT VLLTMLSSAERGCPKGCRCEGKMVYCESQKLQEIPSSISAGCLG  
LSLRYNLSLQKLKYNQFKGLNQLTWLYLDHNNHISNIDENAFNGIRRLKELILSSNRISYFLNNTFR  
PVTNLRNLDLSYNQLHSLGSEQFRGLRKLLSLHLRSNSLRTIPVRI FQDCRNLELLDLGYNRIRS  
LARNVFAGMIRLKEHLEHNQFSKLNALFPRLVSLQONLYLQWNKISVIGQTMSTWSSLQRLDL  
SGNEIEAFSGPSVFQCVPNLQRLNLD SNKLTFIGQEILDSWISLNDISLAGNIWECSRNICSLVN  
WLKSFKGLRENTIICASPKELOGVNVIDAVKNYSICGKSTTERFDLARALPKPTFKPKLPRPKHE  
SKPPLPPTVGATEPGPETDADAEHISFHKIIAGSVALFLSVLVILLVIYVSWKRYPASMKQLQQR  
SLMRRHRKKKRQSLKQMT PSTQEFYVDYKPTNTETSEM LLLNGTGPCTYNKSGSRECEV

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-33

#### **Transmembrane domain:**

amino acids 420-442

#### **N-glycosylation sites.**

amino acids 126-129, 357-360, 496-499, 504-507

#### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 465-468

#### **Tyrosine kinase phosphorylation site.**

amino acids 136-142

#### **N-myristoylation sites.**

amino acids 11-16, 33-38, 245-250, 332-337, 497-502, 507-512

**FIGURE 125**

CCGTTATCGTCTTGCGCTACTGCTGA**ATG**TCCGTCCCGGAGGAGGAGGAGAGGCTTTTGCCGCTG  
ACCCAGAGATGGCCCCGAGCGAGCAAATTCCCTACTGTCCGGCTGCGCGGCTACCGTGGCCGAGCT  
AGCAACCTTTCCCCTGGATCTCACAAAACTCGACTCCAAATGCAAGGAGAAGCAGCTCTTGCTC  
GGTTGGGAGACGGTGCAAGAGAATCTGCCCCCTATAGGGGAATGGTGCGCACAGCCCTAGGGATC  
ATTGAAGAGGAAGGCTTTCTAAAGCTTTGGCAAGGAGTGACACCCGCCATTTACAGACACGTAGT  
GTATTCTGGAGGTCGAATGGTCACATATGAACATCTCCGAGAGGTTGTGTTTGGCAAAAGTGAAG  
ATGAGCATTATCCCCTTTGGAAATCAGTCATTGGAGGGATGATGGCTGGTGTTATTGGCCAGTTT  
TTAGCCAATCCAACCTGACCTAGTGAAGGTTCCAGATGCAAATGGAAGGAAAAGGAAACTGGAAGG  
AAAACCATTTGCGATTTTCGTGGTGTACATCATGCATTTGCAAAAATCTTAGCTGAAGGAGGAATAC  
GAGGGCTTTGGGCAGGCTGGGTACCCAATATACAAAGAGCAGCACTGGTGAATATGGGAGATTTA  
ACCACTTATGATACAGTGAAACACTACTTGGTATTGAATACACCACTTGAGGACAATATCATGAC  
TCACGGTTTATCAAGTTTATGTTCTGGACTGGTAGCTTCTATTCTGGGAACACCAGCCGATGTCA  
TCAAAGCAGAATAATGAATCAACCACGAGATAAACAAGGAAGGGGACTTTTGTATAAATCATCG  
ACTGACTGCTTGATTCAAGGCTGTTCAAGGTGAAGGATTCATGAGTCTATATAAAGGCTTTTTACC  
ATCTTGGCTGAGAATGACCCCTTGGTCAATGGTGTCTGGCTTACTTATGAAAAATCAGAGAGA  
TGAGTGGAGTCAGTCCATTT**TAA**

## **FIGURE 126**

MSVPEEEERLLPLTQRWPRASKFLLSGCAATVAELATFPLDLTKTRLQMQGEAALARLGDGARES  
APYRGMVRTALGIIEEEGFLKLWQGVTPAIYRHVVYSGGRMVITYEHLREVVFVGKSEDEHYPLWKS  
VIGGMMAGVIGQFLANPTDLVKVQMOMEGKRKLEGKPLRFRGVHHAFAKILAEGGIRGLWAGWVP  
NIQRAALVNMGDLTTYDTVKHYLVNTPLEDNIMTHGLSSLCSGLVASILGTPADVIKSRIMNQP  
RDKQGRGLLYKSSTDCLIQAVQGEGFMSLYKGFLPSWLRMTPWSMVFWLTYEKIREMSGVSPF

**Transmembrane domains:**

amino acids 25-38, 130-147, 233-248



**FIGURE 127**

CGCGGATCGGACCCAAGCAGGTCGGCGGCGGGCGGCAGGAGAGCGGCCGGGCGTCAGCTCCTCGAC  
CCCCGTGTCGGGCTAGTCCAGCGAGGCGGACGGGCGGCGTGGGCCCATGGCCAGGCCCGGCATGG  
AGCGGTGGCGCGACCGGCTGGCGCTGGTGACGGGGGCTCGGGGGGCATCGGCGCGGCCGTGGCC  
CGGGCCCTGGTCCAGCAGGGA<sup>^</sup>CTGAAGGTGGTGGGCTGCGCCCGCACTGTGGGCAACATCGAGGA  
GCTGGCTGCTGAATGTAAGAGTGCAGGCTACCCCGGGACTTTGATCCCCTACAGATGTGACCTAT  
CAAATGAAGAGGACATCCTCTCCATGTTCTCAGCTATCCGTTCTCAGCACAGCGGTGTAGACATC  
TGCATCAACAATGCTGGCTTGGCCCCGGCCTGACACCCTGCTCTCAGGCAGCACCAAGTGGTTGGAA  
GGACATGTTCAATGTGAACGTGCTGGCCCTCAGCATCTGCACACGGGAAGCCTACCAGTCCATGA  
AGGAGCGGAATGTGGACGATGGGCACATCATTAACATCAATAGCATGTCTGGCCACCGAGTGTTA  
CCCCTGTCTGTGACCCACTTCTATAGTGCCACCAAGTATGCCGTCACTGCGCTGACAGAGGGACT  
GAGGCAAGAGCTTCGGGAGGCCCAGACCCACATCCGAGCCACGTGCATCTCTCCAGGTGTGGTGG  
AGACACAATTGCGCTTCAAACCTCCACGACAAGGACCCTGAGAAGGCAGCTGCCACCTATGAGCAA  
ATGAAGTGTCTCAAACCCGAGGATGTGGCCGAGGCTGTTATCTACGTCCTCAGCACCCCCGCACA  
CATCCAGATTGGAGACATCCAGATGAGGCCACGGAGCAGGTGACCTAGTGACTGTGGGAGCTCC  
TCCTTCCCTCCCCACCCTTCATGGCTTGCCTCCTGCCTCTGGATTTTAGGTGTTGATTTCTGGAT  
CACGGGATACCACTTCCTGTCCACACCCCGACCAGGGGCTAGAAAATTTGTTTGAGATTTTATA  
TCATCTTGTCAAATTGCTTCAGTTGTAAATGTGAAAAATGGGCTGGGGAAAGGAGGTGGTGTCCC  
TAATTGTTTTACTTGTTAACTTGTTCTTGTCGCCCTGGGCACTTGGCCTTTGTCTGCTCTCAGTG  
TCTTCCCTTTGACATGGGAAAGGAGTTGTGGCCAAAATCCCCATCTTCTTGACCTCAACGTCTG  
TGGCTCAGGGCTGGGGTGGCAGAGGGAGGCCTTCACCTTATATCTGTGTTGTTATCCAGGGCTCC  
AGACTTCCTCCTCTGCCTGCCCCACTGCACCCTCTCCCCCTTATCTATCTCCTTCTCGGCTCCCC  
AGCCCAGTCTTGGCTTCTTGTCCCCTCCTGGGGTCATCCCTCCACTCTGACTCTGACTATGGCAG  
CAGAACACCAGGGCCTGGCCCAGTGGATTTTCATGGTGATCATTAAGAAAGAAAAATCGCAACCAA  
AAAAAAAAAA

## **FIGURE 128**

MARPGMERWRDRLALVTGASGGIGA AAVARALVQQGLKVVG CARTVGNIEELAAECKSAGYPGTLI  
PYRCDLSNEEDILSMFSAIRSQHSGVDICINNAGLARPD TLLSGSTSGWKDMFNVNVLALSICTR  
EAYQSMKERNVDDGHIININMSGHRVLP LSVTHFY SATKYAVTALTEGLRQELREAQTHIRATC  
ISPGVVETQFAFKLHDKDPEKAAATYEQMKCLKPEDVAEAVIYVLSTPAHIQIGDIQMRPTEQVT

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-17

#### **N-myristoylation sites.**

amino acids 18-24, 21-27, 22-28, 24-30, 40-46, 90-96, 109-115, 199-205

#### **Short-chain alcohol dehydrogenase.**

amino acids 30-42, 104-114

**FIGURE 129**

AACTTCTACATGGGCCTCCTGCTGCTGGTGCTCTTCCTCAGCCTCCTGCCGGTGGCCTACACCAT  
CATGTCCCTCCCACCCTCCTTTGACTGCGGGCCGTTCAGGTGCAGAGTCTCAGTTGCCCGGGAGC  
ACCTCCCCCTCCCGAGGCAGTCTGCTCAGAGGGCCTCGGCCCAGAATTCCAGTTCTGGTTTCATGC  
CAGCCTGTAAAAGGCCATGGAACCTTTGGGTGAATCACCGATGCCATTTAAGAGGGTTTTCTGCCA  
GGATGGAAATGTTAGGTCGTTCTGTGTCTGCGCTGTTCAATTCAGTAGCCACCAGCCACCTGTGG  
CCGTTGAGTGCTTGAAATGAGGAACTGAGAAAATTAATTTCTCATGTATTTTTCTCATTTATTTA  
TTAATTTTTTAAGTATAGTTGTACATATTTGGGGGTACATGTGATATTTGGATACATGTATACAA  
TATATAATGATCAAATCAGGGTAACTGGGATATCCATCACATCAAACATTTATTTTTTATTCTTT  
TTAGACAGAGTCTCACTCTGTCACCCAGGCTGGAGTGCAGTGGTGCCATCTCAGCTTACTGCAAC  
CTCTGCCTGCCAGGTTCAAGCGATTCTCATGCCTCCACCTCCCAAGTAGCTGGGACTACAGGCAT  
GCACCACAATGCCCAACTAATTTTTGTATTTTTTAGTAGAGACGGGGTTTTGCCATGTTGCCCAGG  
CTGGCCTTGAACTCCTGGCCTCAAACAATCCACTTGCCTCGGCCTCCCAAAGTGTTATGATTACA  
GGCGTGAGCCACCGTGCCTGGCCTAAACATTTATCTTTTCTTTGTGTTGGGAACCTTGAAATTAT  
ACAATGAATTATTGTAACTGTCATCTCCCTGCTGTGCTATGGAACACTGGGACTTCTTCCCTCT  
ATCTAACTGTATATTTGTACCAGTTAACCAACCGTACTTCATCCCCACTCCTCTCTATCCTTCCC  
AACCTCTGATCACCTCATTCTACTCTCTACCTCCATGAGATCCACTTTTTTAGCTCCCACATGTG  
AGTAAGAAAATGCAATATTTGTCTTTCTGTGCCTGGCTTATTTCACTTAACATAATGACTTCCTG  
TTCCATCCATGTTGCTGCAAATGACAGGATTTTCGTTCTTAATTTCAATTAAAATAACCACACATG  
GCAAAAA

130/168

## **FIGURE 130**

MGLLLLVLFLSLLPVAYTIMSLPPSFDCGPFRCRVSVAREHLPSRGSLLRGPRPRI PVLVSCQPV  
KGHGTLGESPMPFKRVFQCQDGNVRSFCVCAVHFSSHQPPVAVECLK

**Important features of the protein:**

**Signal peptide:**

amino acids 1-18

**N-myristoylation site.**

amino acids 86-92

**Zinc carboxypeptidases, zinc-binding region 2 signature.**

amino acids 68-79

**FIGURE 131**

TTCTGAAGTAACGGAAGCTACCTTGTATAAAGACCTCAACACTGCTGACCATGATCAGCGCAGCC  
TGGAGCATCTTCCTCATCGGGACTAAAATTGGGCTGTTCCCTTCAAGTAGCACCTCTATCAGTTAT  
GGCTAAATCCTGTCCATCTGTGTGTGCTGCGATGCGGGTTTCATTTACTGTAATGATCGCTTTC  
TGACATCCATTCCAACAGGAATACCAGAGGATGCTACAACCTCTCTACCTTCAGAACAAACAAATA  
AATAATGCTGGGATTCCCTTCAGATTTGAAAACTTGCTGAAAGTAGAAAGAATATACCTATACCA  
CAACAGTTTAGATGAATTTCCCTACCAACCTCCCAAAGTATGTAAAAGAGTTACATTTGCAAGAAA  
ATAACATAAGGACTATCACTTATGATTCACTTTCAAAAATTCCTTATCTGGAAGAATTACATTTA  
GATGACAACTCTGTCTCTGCAAGTAGCATAGAAGAGGGAGCATTCGAGACAGCAACTATCTCCG  
ACTGCTTTTCCCTGTCCCGTAATCACCTTAGCACAATTCCTTGGGGTTTGCCCAAGGACTATAGAAG  
AACTACGCTTGGAATGATAATCGCATATCCACTATTTTCATCACCATCTCTTCAAGGTCTCACTAGT  
CTAAAACGCCTGGTTCTAGATGGAAACCTGTTGAACAATCATGGTTTAGGTGACAAAGTTTTCTT  
CAACCTAGTTAATTTGACAGAGCTGTCCCTGGTGCGGAATTCCTGACTGCTGCACCAGTAAACC  
TTCCAGGCACAAACCTGAGGAAGCTTTATCTTCAAGATAACCACATCAATCGGGTGCCCCCAAAT  
GCTTTTTCTTATCTAAGGCAGCTCTATCGACTGGATATGTCCAATAATAACCTAAGTAATTTACC  
TCAGGGTATCTTTGATGATTTGGACAATATAACACAACCTGATTCTTCGCAACAATCCCTGGTATT  
GCGGGTGCAAGATGAAATGGGTACGTGACTGGTTACAATCACTACCTGTGAAGGTCAACGTGCGT  
GGGCTCATGTGCCAAGCCCCAGAAAAGGTTTCGTGGGATGGCTATTAAGGATCTCAATGCAGAACT  
GTTTGATTGTAAGGACAGTGGGATTGTAAGCACCATTTCAGATAACCACTGCAATACCCAACACAG  
TGTATCCTGCCAAGGACAGTGGCCAGCTCCAGTGACCAAACAGCCAGATATTAAGAACCCCAAG  
CTCACTAAGGATCAACAAACCACAGGGAGTCCCTCAAGAAAAACAATTACAATTACTGTGAAGTC  
TGTCACCTCTGATACCATTTCATATCTCTTGGAACCTTGCTCTACCTATGACTGCTTTGAGACTCA  
GCTGGCTTAAACTGGGCCATAGCCCGGCATTTGGATCTATAACAGAAACAATTGTAACAGGGGAA  
CGCAGTGAGTACTTGGTCACAGCCCTGGAGCCTGATTACCCCTATAAAGTATGCATGGTTCCCAT  
GGAAACCAGCAACCTCTACCTATTTGATGAAACTCCTGTTTGTATTGAGACTGAAACTGCACCCC  
TTCGAATGTACAACCCTACAACCACCCTCAATCGAGAGCAAGAGAAAGAACCTTACAAAAACCCC  
AATTTACCTTTGGCTGCCATCATTGGTGGGGCTGTGGCCCTGGTTACCATTGCCCTTCTTGCTTT  
AGTGTGTTGGTATGTTTCATAGGAATGGATCGCTCTTCTCAAGGAACTGTGCATATAGCAAAGGGA  
GGAGAAGAAAGGATGACTATGCAGAAGCTGGCACTAAGAAGGACAACTCTATCCTGGAAATCAGG  
GAACTTCTTTTTCAGATGTTACCAATAAGCAATGAACCCATCTCGAAGGAGGAGTTTGTAATACA  
CACCATATTTCCCTCCTAATGGAATGAATCTGTACAAAAACAATCACAGTGAAAGCAGTAGTAACC  
GAAGCTACAGAGACAGTGGTATTCCAGACTCAGATCACTCACACTCATGATGCTGAAGGACTCAC  
AGCAGACTTGTGTTTTGGGTTTTTTTAAACCTAAGGGAGGTGATGGT

## **FIGURE 132**

MISAAWSIFLIGTKIGLFLQVAPLSVMAKSCPSVCRCDAGFIYCNDRFLTSTIPTGIPEDATTLYL  
 QNNQINNAGIPSDLKNLLKVERIYLYHNSLDEFPTNLPKYVKELHLQENNIRTITYDSLSKIPYL  
 EELHLDDNSVSAVSIEEGAFRDSNYLRLLFLSRNHLSTIPWGLPRTIEELRLDDNRISTISSPSL  
 QGLTSLKRLVLDGNLLNNHGLGDKVFFNLVNLTELSLVRNSLTAAPVNLPGTNLRKLYLQDNHIN  
 RVPPNAFSYLRQLYRLDMSNNNLSNLPQGI FDDLDNITQLILRNNPWYCGCKMKWVRDWLQSLPV  
 KVNVRGLMCQAPEKVRGMAIKDLNAELFDCKDSGIVSTIQITTAIPNTVYPAQGQWPAPVTKQPD.  
 IKNPKLTKDQQTGSPSRKTITITVKSVTSDTIHISWKLALPMTALRLSWLKLGHSPAFGSITET  
 IVTGERSEYLVTALEPDSPYKVCMPMETSNNLYLFDETPVCIETETAPLRMYNPTTTLNREQEKE  
 PYKNPNLPLAAIIGGAVALVTIALALLVCWYVHRNGSLFSRNCAYSKGRRRKDDYAEAGTKKDNS  
 ILEIRETSFQMLPISNEPISKEEFVIHTIFPPNGMNLYKNNHSESSSNRSYRDSGIPDSDHSHS

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-28

#### **Transmembrane domain:**

amino acids 531-552

#### **N-glycosylation sites.**

amino acids 226-229, 282-285, 296-299, 555-558, 626-629, 633-636

#### **Tyrosine kinase phosphorylation site.**

amino acids 515-522

#### **N-myristoylation sites.**

amino acids 12-17, 172-177, 208-213, 359-364, 534-539, 556-561,  
 640-645

#### **Amidation site.**

amino acids 567-570

#### **Leucine zipper pattern.**

amino acids 159-180

#### **Phospholipase A2 aspartic acid active site.**

amino acids 34-44

**FIGURE 133**

CCGTCATCCCCCTGCAGCCACCCTTCCCAGAGTCCTTTGCCCAGGCCACCCCAGGCTTCTTGGA  
GCCCTGCCGGGGCACTTGTCTT**CATG**TCTGCCAGGGGGAGGTGGGAAGGAGGTGGGAGGAGGGCG  
TGCAGAGGCAGTCTGGGCTTGGCCAGAGCTCAGGGTGCTGAGCGTGTGACCAGCAGTGAGCAGAG  
CCCGGCCATGGCCAGCCTGGGGCTGCTGCTCCTGCTCTTACTGACAGCACTGCCACCGCTGTGGT  
CCTCCTCACTGCCTGGGCTGGACACTGCTGAAAGTAAAGCCACCATTGCAGACCTGATCCTGTCT  
GCGCTGGAGAGAGCCACCGTCTTCCTAGAACAGAGGCTGCCTGAAATCAACCTGGATGGCATGGT  
GGGGGTCCGAGTGCTGGAAGAGCAGCTAAAAAGTGTCCGGGAGAAGTGGGCCCAGGAGCCCCCTGC  
TGCAGCCGCTGAGCCTGCGCGTGGGGATGCTGGGGGAGAAGCTGGAGGCTGCCATCCAGAGATCC  
CTCCACTACCTCAAGCTGAGTGATCCCAAGTACCTAAGAGAGTTCCAGCTGACCCTCCAGCCCGG  
GTTTTGGAAGCTCCCACATGCCTGGATCCACACTGATGCCTCCTTGGTGTACCCACGTTCCGGG  
CCCAGGACTCATTCTCAGAGGAGAGAAGTGACGTGTGCCTGGTGCAGCTGCTGGGAACCGGGACG  
GACAGCAGCGAGCCCTGCGGCCTCTCAGACCTCTGCAGGAGCCTCATGACCAAGCCCGGCTGCTC  
AGGCTACTGCCTGTCCCACTGCTCTTCTCCTCTGGGCCAGAATGAGGGGATGCACACAGG  
GACCACTCCAACAGAGCCAGGACTATATCAACCTCTTCTGCGCCAACATGATGGACTTGAACCGC  
AGAGCTGAGGCCATCGGATACGCTACCCCTACCCGGGACATCTTCATGGAACATCATGTTCTG  
TGGAATGGGCGGCTTCTCCGACTTCTACAAGCTCCGGTGGCTGGAGGCCATTCTCAGCTGGCAGA  
AACAGCAGGAAGGATGCTTCGGGGAGCCTGATGCTGAAGATGAAGAATTATCTAAAGCTATTCAA  
TATCAGCAGCATTTTTTCGAGGAGAGTGAAGAGGCGAGAAAAACAATTTCCAGATTCTCGCTCTGT  
TGCTCAGGCTGGAGTACAGTGGCGCAATCTCGGCTCACTGCAACCTTTGCCTCCTGGGTCAAGC  
AATTCTCTTGCTCATCTCCCGAGTAGCTGGGACTACAGGAGCGTGCCACCATACTGGCTAAT  
TTTTATATTTTTTTAGTAGAGACAGGGTTTCATCATGTTGCTCATGCTGGTCTCGAACTCCTGAT  
CTCAAGAGATCCGCCACCTCAGGCTCCCAAAGTGTGGGATTAT**TAG**GTGTGAGCCACCGTGTCTG  
GCTGAAAAGCACTTTCAAAGAGACTGTGTTGAATAAAGGGCCAAGGTTCTTGCCACCCAGCACTC  
ATGGGGGCTCTCTCCCCTAGATGGCTGCTCCTCCCACAACACAGCCACAGCAGTGGCAGCCCTGG  
GTGGCTTCTTATACATCCTGGCAGAATACCCCCCAGCAAACAGAGAGCCACACCCATCCACACCG  
CCACCACCAAGCAGCCGCTGAGACGGACGGTTCCATGCCAGCTGCCTGGAGGAGGAACAGACCCC  
TTTAGTCCTCATCCCTTAGATCCTGGAGGGCACGGATCACATCCTGGGAAGAAGGCATCTGGAGG  
ATAAGCAAAGCCACCCCGACACCCAATCTTGGAAGCCCTGAGTAGGCAGGGCCAGGGTAGGTGGG  
GGCCGGGAGGGACCCAGGTGTGAACGGATGAATAAAGTTCAACTGCAACTGAAAAAAAAAAAA

## **FIGURE 134**

MSARGRWEGGGRRACRGSGLLARAQGAERVTSSEQRPAMASLGLLLLLLLLLTALPPLWSSSLPGLD  
TAESKATIADLILSALERATVFLEQRLPEINLDGMVGVRVLEEQLKSVREKWAQEPLLQPLSLRV  
GMLGEKLEAAIQRSLHYLKLSDPKYLREFQLTLQPGFWKLPHAWIHTDASLVYPTFGPQDSFSEE  
RSDVCLVQLLGTGTDSSSEPCGLSDLCRSLMTKPGCSGYCLSHQLLFFLWARMRGCTQGPLQQSQD  
YINLFCANMMDLNRRAEAIGYAYPTRDIFMENIMFCGMGGFSDFYKLRWLEAILSQKQOEGCFG  
EPDAEDEELSKAIQYQQHFSSRRVKRREKQFPDSRSVAQAGVQWRNLGSLQPLPPGFKQFSCILIP  
SSWDYRSVPPYLANFYIFLVETGFHHVAHAGLELLISRDPPTSGSQSVGL

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-26

#### **Transmembrane domain:**

amino acids 39-56

#### **Tyrosine kinase phosphorylation sites.**

amino acids 149-156, 274-282

#### **N-myristoylation sites.**

amino acids 10-16, 20-26, 63-69, 208-214

#### **Amidation site.**

amino acids 10-14

#### **Glycoprotein hormones beta chain signature 1.**

amino acids 230-237



**EXHIBIT 1**

**MOLECULAR BIOLOGY OF  
THE CELL  
THIRD EDITION**

Text Editor: Miranda Robertson  
Managing Editor: Ruth Adams  
Illustrator: Nigel Orme  
Molecular Model Drawings: Kate Hesketh-Moore  
Director of Electronic Publishing: John M-Roblin  
Computer Specialist: Chuck Bartelt  
Disk Preparation: Carol Winter  
Copy Editor: Shirley M. Cobert  
Production Editor: Douglas Goertzen  
Production Coordinator: Perry Bessas  
Indexer: Maija Hinkle

*Bruce Alberts* received his Ph.D. from Harvard University and is currently President of the National Academy of Sciences and Professor of Biochemistry and Biophysics at the University of California, San Francisco. *Dennis Bray* received his Ph.D. from the Massachusetts Institute of Technology and is currently a Medical Research Council Fellow in the Department of Zoology, University of Cambridge. *Julian Lewis* received his D.Phil. from the University of Oxford and is currently a Senior Scientist in the Imperial Cancer Research Fund Developmental Biology Unit, University of Oxford. *Martin Raff* received his M.D. from McGill University and is currently a Professor in the MRC Laboratory for Molecular Cell Biology and the Biology Department, University College London. *Keith Roberts* received his Ph.D. from the University of Cambridge and is currently Head of the Department of Cell Biology, the John Innes Institute, Norwich. *James D. Watson* received his Ph.D. from Indiana University and is currently Director of the Cold Spring Harbor Laboratory. He is the author of *Molecular Biology of the Gene* and, with Francis Crick and Maurice Wilkins, won the Nobel Prize in Medicine and Physiology in 1962.

© 1983, 1989, 1994 by Bruce Alberts, Dennis Bray, Julian Lewis, Martin Raff, Keith Roberts, and James D. Watson.

All rights reserved. No part of this book covered by the copyright hereon may be reproduced or used in any form or by any means—graphic, electronic, or mechanical, including photocopying, recording, taping, or information storage and retrieval systems—without permission of the publisher.

**Library of Congress Cataloging-in-Publication Data**

Molecular biology of the cell / Bruce Alberts . . . [et al.].—3rd ed.  
p. cm.

Includes bibliographical references and index.

ISBN 0-8153-1619-4 (hard cover).—ISBN 0-8153-1620-8 (pbk.)

1. Cytology. 2. Molecular biology. I. Alberts, Bruce.

[DNLM: 1. Cells. 2. Molecular Biology. QH 581.2 M718 1994]

QH581.2.M64 1994

574.87—dc20

DNLM/DLC

for Library of Congress

93-45907

CIP

Published by Garland Publishing, Inc.  
717 Fifth Avenue, New York, NY 10022

Printed in the United States of America

15 14 13 12 10 9 8 7

**Front cover:** The photograph shows a rat nerve cell in culture. It is labeled ( *yellow* ) with a fluorescent antibody that stains its cell body and dendritic processes. Nerve terminals ( *green* ) from other neurons (not visible), which have made synapses on the cell, are labeled with a different antibody. (Courtesy of Olaf Mundigl and Pietro de Camilli.)

**Dedication page:** Gavin Borden, late president of Garland Publishing, weathered in during his mid-1980s climb near Mount McKinley with MBoC author Bruce Alberts and famous mountaineer guide Mugs Stump (1940–1992).

**Back cover:** The authors, in alphabetical order, crossing Abbey Road in London on their way to lunch. Much of this third edition was written in a house just around the corner. (Photograph by Richard Olivier.)

extracts. If these minor cell proteins differ among cells to the same extent as the more abundant proteins, as is commonly assumed, only a small number of protein differences (perhaps several hundred) suffice to create very large differences in cell morphology and behavior.

### A Cell Can Change the Expression of Its Genes in Response to External Signals<sup>3</sup>

Most of the specialized cells in a multicellular organism are capable of altering their patterns of gene expression in response to extracellular cues. If a liver cell is exposed to a glucocorticoid hormone, for example, the production of several specific proteins is dramatically increased. Glucocorticoids are released during periods of starvation or intense exercise and signal the liver to increase the production of glucose from amino acids and other small molecules; the set of proteins whose production is induced includes enzymes such as tyrosine aminotransferase, which helps to convert tyrosine to glucose. When the hormone is no longer present, the production of these proteins drops to its normal level.

Other cell types respond to glucocorticoids in different ways. In fat cells, for example, the production of tyrosine aminotransferase is reduced, while some other cell types do not respond to glucocorticoids at all. These examples illustrate a general feature of cell specialization—different cell types often respond in different ways to the same extracellular signal. Underlying this specialization are features that do not change, which give each cell type its permanently distinctive character. These features reflect the persistent expression of different sets of genes.

### Gene Expression Can Be Regulated at Many of the Steps in the Pathway from DNA to RNA to Protein<sup>4</sup>

If differences between the various cell types of an organism depend on the particular genes that the cells express, at what level is the control of gene expression exercised? There are many steps in the pathway leading from DNA to protein, and all of them can in principle be regulated. Thus a cell can control the proteins it makes by (1) controlling when and how often a given gene is transcribed (**transcriptional control**), (2) controlling how the primary RNA transcript is spliced or otherwise processed (**RNA processing control**), (3) selecting which completed mRNAs in the cell nucleus are exported to the cytoplasm (**RNA transport control**), (4) selecting which mRNAs in the cytoplasm are translated by ribosomes (**translational control**), (5) selectively destabilizing certain mRNA molecules in the cytoplasm (**mRNA degradation control**), or (6) selectively activating, inactivating, or compartmentalizing specific protein molecules after they have been made (**protein activity control**) (Figure 9-2).

For most genes transcriptional controls are paramount. This makes sense because, of all the possible control points illustrated in Figure 9-2, only transcriptional control ensures that no superfluous intermediates are synthesized. In the

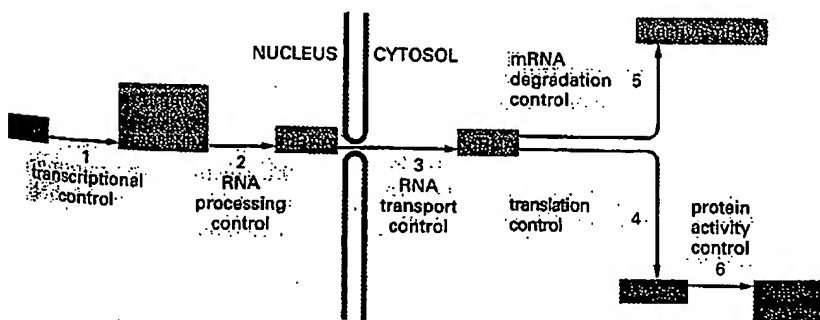


Figure 9-2 Six steps at which eucaryote gene expression can be controlled. Only controls that operate at steps 1 through 5 are discussed in this chapter. The regulation of protein activity (step 6) is discussed in Chapter 5; this includes reversible activation or inactivation by protein phosphorylation as well as irreversible inactivation by proteolytic degradation.

following sections we discuss the DNA and protein components that regulate the initiation of gene transcription. We return at the end of the chapter to the other ways of regulating gene expression.

## Summary

*The genome of a cell contains in its DNA sequence the information to make many thousands of different protein and RNA molecules. A cell typically expresses only a fraction of its genes, and the different types of cells in multicellular organisms arise because different sets of genes are expressed. Moreover, cells can change the pattern of genes they express in response to changes in their environment, such as signals from other cells. Although all of the steps involved in expressing a gene can in principle be regulated, for most genes the initiation of RNA transcription is the most important point of control.*

## DNA-binding Motifs in Gene Regulatory Proteins<sup>5</sup>

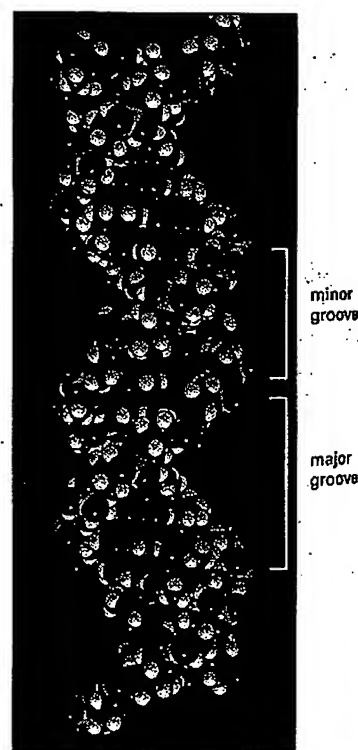
How does a cell determine which of its thousands of genes to transcribe? As discussed in Chapter 8, the transcription of each gene is controlled by a regulatory region of DNA near the site where transcription begins. Some regulatory regions are simple and act as switches that are thrown by a single signal. Other regulatory regions are complex and act as tiny microprocessors, responding to a variety of signals that they interpret and integrate to switch the neighboring gene on or off. Whether complex or simple, these switching devices consist of two fundamental types of components: (1) short stretches of DNA of defined sequence and (2) *gene regulatory proteins* that recognize and bind to them.

We begin our discussion of gene regulatory proteins by describing how these proteins were discovered.

## Gene Regulatory Proteins Were Discovered Using Bacterial Genetics<sup>6</sup>

Genetic analyses in bacteria carried out in the 1950s provided the first evidence of the existence of **gene regulatory proteins** that turn specific sets of genes on or off. One of these regulators, the *lambda repressor*, is encoded by a bacterial virus, *bacteriophage lambda*. The repressor shuts off the viral genes that code for the protein components of new virus particles and thereby enables the viral genome to remain a silent passenger in the bacterial chromosome, multiplying with the bacterium when conditions are favorable for bacterial growth (see Figure 6-80). The lambda repressor was among the first gene regulatory proteins to be characterized, and it remains one of the best understood, as we discuss later. Other bacterial regulators respond to nutritional conditions by shutting off genes encoding specific sets of metabolic enzymes when they are not needed. The *lac repressor*, for example, the first of these bacterial proteins to be recognized, turns off the production of the proteins responsible for lactose metabolism when this sugar is absent from the medium.

The first step toward understanding gene regulation was the isolation of mutant strains of bacteria and bacteriophage lambda that were unable to shut off specific sets of genes. It was proposed at the time, and later proved, that most of these mutants were deficient in proteins acting as specific repressors for these sets of genes. Because these proteins, like most gene regulatory proteins, are present in small quantities, it was difficult and time-consuming to isolate them. They were eventually purified by fractionating cell extracts on a series of standard chromatography columns (see pp. 166-169). Once isolated, the proteins were shown to bind to specific DNA sequences close to the genes that they



**Figure 9-3 Double-helical structure of DNA.** The major and minor grooves on the outside of the double helix are indicated. The atoms are colored as follows: carbon, dark blue; nitrogen, light blue; hydrogen, white; oxygen, red; phosphorus, yellow.

regulate  
by a cor  
experim

The O

As disci  
double l  
otide se  
these pl  
pairs in  
and anc  
ded wit  
without  
at the s  
bond d  
to reco  
major g  
(Figure  
for gro  
Alt  
most ir  
only or  
double

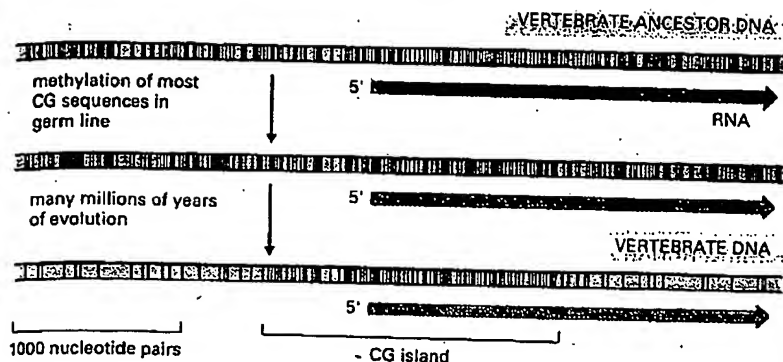


Figure 9-71 A mechanism to explain both the marked deficiency of CG sequences and the presence of CG islands in vertebrate genomes. A black line marks the location of an unmethylated CG dinucleotide in the DNA sequence, while a red line marks the location of a methylated CG dinucleotide.

## Summary

The many types of cells in animals and plants are created largely through mechanisms that cause different genes to be transcribed in different cells. Since many specialized animal cells can maintain their unique character when grown in culture, the gene regulatory mechanisms involved in creating them must be stable once established and heritable when the cell divides, endowing the cell with a memory of its developmental history. Prokaryotes and yeasts provide unusually accessible model systems in which to study gene regulatory mechanisms, some of which may be relevant to the creation of specialized cell types in higher eucaryotes. One such mechanism involves a competitive interaction between two (or more) gene regulatory proteins, each of which inhibits the synthesis of the other; this can create a flip-flop switch that switches a cell between two alternative patterns of gene expression. Direct or indirect positive feedback loops, which enable gene regulatory proteins to perpetuate their own synthesis, provide a general mechanism for cell memory.

In eucaryotes gene transcription is generally controlled by combinations of gene regulatory proteins. It is thought that each type of cell in a higher eucaryotic organism contains a specific combination of gene regulatory proteins that ensures the expression of only those genes appropriate to that type of cell. A given gene regulatory protein may be expressed in a variety of circumstances and typically is involved in the regulation of many genes.

In addition to diffusible gene regulatory proteins, inherited states of chromatin condensation are also utilized by eucaryotic cells to regulate gene expression. In vertebrates DNA methylation also plays a part, mainly as a device to reinforce decisions about gene expression that are made initially by other mechanisms.

## Posttranscriptional Controls

Although controls on the initiation of gene transcription are the predominant form of regulation for most genes, other controls can act later in the pathway from RNA to protein to modulate the amount of gene product that is made. Although these posttranscriptional controls, which operate after RNA polymerase has bound to the gene's promoter and begun RNA synthesis, are less common than transcriptional control, for many genes they are crucial. It seems that every step in gene expression that could be controlled in principle is likely to be regulated under some circumstances for some genes.

We consider the varieties of posttranscriptional regulation in temporal order, according to the sequence of events that might be experienced by an RNA molecule after its transcription has begun (Figure 9-72).

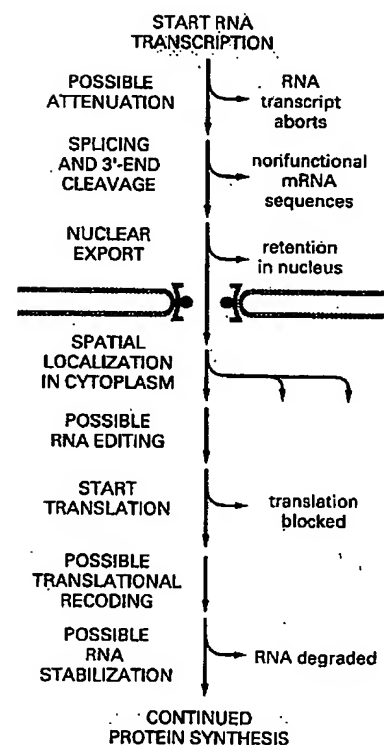


Figure 9-72 Possible posttranscriptional controls on gene expression. Only a few of these controls are likely to be used for any one gene.

# CHAPTER 29

## Regulation of transcription

Genes VII (1997) CH 29, pp. 847-848.  
Benjamin Lewin

The phenotypic differences that distinguish the various kinds of cells in a higher eukaryote are largely due to differences in the expression of genes that code for proteins, that is, those transcribed by RNA polymerase II. In principle, the expression of these genes might be regulated at any one of several stages. The concept of the "level of control" implies that gene expression is not necessarily an automatic process once it has begun. It could be regulated in a gene-specific way at any one of several sequential steps. We can distinguish (at least) five potential control points, forming the series:

Activation of gene structure  
↓  
Initiation of transcription  
↓  
Processing the transcript  
↓  
Transport to cytoplasm  
↓  
Translation of mRNA

The existence of the first step is implied by the discovery that genes may exist in either of two structural conditions. Relative to the state of most of the genome, genes are found in an "active" state in the cells in which they are expressed (see Chapter 27). The change of structure is distinct from the act of transcription, and indicates that the gene is "transcribable." This suggests that acquisition of the "active" structure must be the first step in gene expression.

Transcription of a gene in the active state is

controlled at the stage of initiation, that is, by the interaction of RNA polymerase with its promoter. This is now becoming susceptible to analysis in the *in vitro* systems (see Chapter 28). For most genes, this is a major control point; probably it is the most common level of regulation.

There is at present no evidence for control at subsequent stages of transcription in eukaryotic cells, for example, via antitermination mechanisms.

The primary transcript is modified by capping at the 5' end, and usually also by polyadenylation at the 3' end. Introns must be spliced out from the transcripts of interrupted genes. The mature RNA must be exported from the nucleus to the cytoplasm. Regulation of gene expression by selection of sequences at the level of nuclear RNA might involve any or all of these stages, but the one for which we have most evidence concerns changes in splicing; some genes are expressed by means of alternative splicing patterns whose regulation controls the type of protein product (see Chapter 30).

Finally, the translation of an mRNA in the cytoplasm can be specifically controlled. There is little evidence for the employment of this mechanism in adult somatic cells, but it does occur in some embryonic situations, as described in Chapter 7. The mechanism is presumed to involve the blocking of initiation of translation of some mRNAs by specific protein factors.

But having acknowledged that control of gene expression can occur at multiple stages, and that production of RNA cannot inevitably be equated with production of protein, it is clear

that the overwhelming majority of regulatory events occur at the initiation of transcription. Regulation of tissue-specific gene transcription lies at the heart of eukaryotic differentiation; indeed, we see examples in Chapter 38 in which proteins that regulate embryonic development prove to be transcription factors. A regulatory transcription factor serves to provide

common control of a large number of target genes, and we seek to answer two questions about this mode of regulation: what identifies the common target genes to the transcription factor; and how is the activity of the transcription factor itself regulated in response to intrinsic or extrinsic signals?

## Response elements identify genes under common regulation

The principle that emerges from characterizing groups of genes under common control is that *they share a promoter element that is recognized by a regulatory transcription factor*. An element that causes a gene to respond to such a factor is called a response element; examples are the HSE (heat shock response element), GRE (glucocorticoid response element), SRE (serum response element).

The properties of some inducible transcription factors and the elements that they recognize are summarized in Table 29.1. Response elements have the same general characteristics as upstream elements of promoters or enhancers. They contain short consensus sequences, and copies of the response elements found in different genes are closely related, but not necessarily identical. The region bound by the factor extends for a short distance on either side of

the consensus sequence. In promoters, the elements are not present at fixed distances from the startpoint, but are usually <200 bp upstream of it. The presence of a single element usually is sufficient to confer the regulatory response, but sometimes there are multiple copies.

Response elements may be located in promoters or in enhancers. Some types of elements are typically found in one rather than the other: usually an HSE is found in a promoter, while a GRE is found in an enhancer. We assume that all response elements function by the same general principle. A gene is regulated by a sequence at the promoter or enhancer that is recognized by a specific protein. The protein functions as a transcription factor needed for RNA polymerase to initiate. Active protein is available only under conditions when the gene is to be expressed; its absence means that the promoter is not activated by this particular circuit.

An example of a situation in which many genes are controlled by a single factor is provided by the heat shock response. This is common to a wide range of prokaryotes and eukaryotes and involves multiple controls of gene expression: an increase in temperature turns off transcription of some genes, turns on transcription of the heat shock genes, and causes changes in the translation of mRNAs. The control of the heat shock genes illustrates the differences between 'prokaryotic' and eukaryotic modes of control. In bacteria, a new sigma factor is synthesized that directs RNA polymerase holoenzyme to recognize an alter-

Table 29.1 Inducible transcription factors bind to response elements that identify groups of promoters or enhancers subject to coordinate control.

Regulatory Agent	Module	Consensus	Factor
Heat shock	HSE	CNNGAANNTCGNG	HSTF
Glucocorticoid	GRE	TGGTACAAATGTTCT	Receptor
Phorbol ester	TRE	TGACTCA	AP1
Serum	SRE	CCATATTAGG	SRF

Research

Open Access

**Prostate stem cell antigen (PSCA) expression in human prostate cancer tissues and its potential role in prostate carcinogenesis and progression of prostate cancer**Zhao Zhigang\*<sup>1</sup> and Shen Wenlv<sup>2</sup>

Address: <sup>1</sup>Department of Urology, Shantou University Medical College, Shantou, Guangdong, China and <sup>2</sup>Department of Urology, No 2. Affiliated Hospital of Shantou University Medical College, Shantou, Guangdong, China

Email: Zhao Zhigang\* - zgzhao@yaho.com.cn; Shen Wenlv - wshen99@hotmail.com

\* Corresponding author

Published: 10 May 2004

Received: 30 March 2004

World Journal of Surgical Oncology 2004, 2:13

Accepted: 10 May 2004

This article is available from: <http://www.wjso.com/content/2/1/13>

© 2004 Zhigang and Wenlv; licensee BioMed Central Ltd. This is an Open Access article: verbatim copying and redistribution of this article are permitted in all media for any purpose, provided this notice is preserved along with the article's original URL.

**Abstract**

**Background:** Prostate stem cell antigen (PSCA) is a recently defined homologue of the Thy-1/Ly-6 family of glycosylphosphatidylinositol (GPI)-anchored cell surface antigens. The purpose of the present study was to examine the expression status of PSCA protein and mRNA in clinical specimens of human prostate cancer (Pca) and to validate it as a potential molecular target for diagnosis and treatment of Pca.

**Materials and Methods:** Immunohistochemical (IHC) and *in situ* hybridization (ISH) analyses of PSCA expression were simultaneously performed on paraffin-embedded sections from 20 benign prostatic hyperplasia (BPH), 20 prostatic intraepithelial neoplasm (PIN) and 48 prostate cancer (Pca) tissues, including 9 androgen-independent prostate cancers. The level of PSCA expression was semiquantitatively scored by assessing both the percentage and intensity of PSCA-positive staining cells in the specimens. Then compared PSCA expression between BPH, PIN and Pca tissues and analysed the correlations of PSCA expression level with pathological grade, clinical stage and progression to androgen-independence in Pca.

**Results:** In BPH and low grade PIN, PSCA protein and mRNA staining were weak or negative and less intense and uniform than that seen in HGPIN and Pca. There were moderate to strong PSCA protein and mRNA expression in 8 of 11 (72.7%) HGPIN and in 40 of 48 (83.4%) Pca specimens examined by IHC and ISH analyses, with statistical significance compared with BPH (20%) and low grade PIN (22.2%) samples ( $p < 0.05$ , respectively). The expression level of PSCA increased with high Gleason grade, advanced stage and progression to androgen-independence ( $p < 0.05$ , respectively). In addition, IHC and ISH staining showed a high degree of correlation between PSCA protein and mRNA overexpression.

**Conclusions:** Our data demonstrate that PSCA as a new cell surface marker is overexpressed by a majority of human Pca. PSCA expression correlates positively with adverse tumor characteristics, such as increasing pathological grade (poor cell differentiation), worsening clinical stage and androgen-independence, and speculatively with prostate carcinogenesis. PSCA protein overexpression results from upregulated transcription of PSCA mRNA. PSCA may have prognostic utility and may be a promising molecular target for diagnosis and treatment of Pca.



## Introduction

Prostate cancer (Pca) is the second leading cause of cancer-related death in American men and is becoming a common cancer increasing in China. Despite recently great progress in the diagnosis and management of localized disease, there continues to be a need for new diagnostic markers that can accurately discriminate between indolent and aggressive variants of Pca. There also continues to be a need for the identification and characterization of potential new therapeutic targets on Pca cells. Current diagnostic and therapeutic modalities for recurrent and metastatic Pca have been limited by a lack of specific target antigens of Pca.

Although a number of prostate-specific genes have been identified (i.e. prostate specific antigen, prostatic acid phosphatase, glandular kallikrein 2), the majority of these are secreted proteins not ideally suited for many immunological strategies. So, the identification of new cell surface antigens is critical to the development of new diagnostic and therapeutic approaches to the management of Pca.

Reiter RE et al [1] reported the identification of prostate stem cell antigen (PSCA), a cell surface antigen that is predominantly prostate specific. The PSCA gene encodes a 123 amino acid glycoprotein, with 30% homology to stem cell antigen 2 (Sca 2). Like Sca-2, PSCA also belongs to a member of the Thy-1/Ly-6 family and is anchored by a glycosylphosphatidylinositol (GPI) linkage. mRNA *in situ* hybridization (ISH) localized PSCA expression in normal prostate to the basal cell epithelium, the putative stem cell compartment of prostatic epithelium, suggesting that PSCA may be a marker of prostate stem/progenitor cells.

In order to examine the status of PSCA protein and mRNA expression in human Pca and validate it as a potential diagnostic and therapeutic target for Pca, we used immunohistochemistry (IHC) and *in situ* hybridization (ISH) simultaneously, and conducted PSCA protein and mRNA expression analyses in paraffin-embedded tissue specimens of benign prostatic hyperplasia (BPH, n = 20), prostate intraepithelial neoplasm (PIN, n = 20) and prostate cancer (Pca, n = 48). Furthermore, we evaluated the possible correlation of PSCA expression level with Pca tumorigenesis, grade, stage and progression to androgen-independence.

## Materials and methods

### Tissue samples

All of the clinical tissue specimens studied herein were obtained from 80 patients of 57–84 years old by prostatectomy, transurethral resection of prostate (TURP) or biopsies. The patients were classified as 20 cases of BPH, 20 cases of PIN, 40 cases of primary Pca, including 9 patients

with recurrent Pca and a history of androgen ablation therapy (orchiectomy and/or hormonal therapy), who were referred to as androgen-independent prostate cancers. Eight specimens were harvested from these androgen-independent Pca patients prior to androgen ablation treatment. Each tissue sample was cut into two parts, one was fixed in 10% formalin for IHC and the other treated with 4% paraformaldehyde/0.1 M PBS PH 7.4 in 0.1% DEPC for 1 h for ISH analysis, and then embedded in paraffin. All paraffin blocks examined were then cut into 5  $\mu$ m sections and mounted on the glass slides specific for IHC and ISH respectively in the usual fashion. H&E-stained section of each Pca was evaluated and assigned a Gleason score by the experienced urological pathologist at our institution based on the criteria of Gleason score [2]. The Gleason sums are summarized in Table 1. Clinical staging was performed according to Jewett-whitmore-prout staging system, as shown in Table 2. In the category of PIN, we graded the specimens into two groups, i.e. low grade PIN (grade I – II) and high grade PIN (HG PIN, grade III) on the basis of literatures [3,4].

### Immunohistochemical (IHC) analysis

Briefly, tissue sections were deparaffinized, dehydrated, and subjected to microwaving in 10 mmol/L citrate buffer, PH 6.0 (Boshide, Wuhan, China) in a 900 W oven for 5 min to induce epitope retrieval. Slides were allowed to cool at room temperature for 30 min. A primary mouse antibody specific to human PSCA (Boshide, Wuhan, China) with a 1:100 dilution was applied to incubate with the slides at room temperature for 2 h. Labeling was detected by sequentially adding biotinylated secondary antibodies and streptavidin-peroxidase, and localized using 3,3'-diaminobenzidine reaction. Sections were then counterstained with hematoxylin. Substitution of the primary antibody with phosphate-buffered-saline (PBS) served as a negative-staining control.

### mRNA *in situ* hybridization (ISH)

Five- $\mu$ m-thick tissue sections were deparaffinized and dehydrated, then digested in pepsin solution (4 mg/ml in 3% citric acid) for 20 min at 37.5°C, and further processed for ISH. Digoxigenin-labeled sense and antisense human PSCA RNA probes (obtained from Boshide, Wuhan, China) were hybridized to the sections at 48°C overnight. The posthybridization wash with a high stringency was performed sequentially at 37°C in 2  $\times$  standard saline citrate (SSC) for 10 min, in 0.5  $\times$  SSC for 15 min and in 0.2  $\times$  SSC for 30 min. The slides were then incubated to biotinylated mouse anti-digoxigenin antibody at 37.5°C for 1 h followed by washing in 1  $\times$  PBS for 20 min at room temperature, and then to streptavidin-peroxidase at 37.5°C for 20 min followed by washing in 1  $\times$  PBS for 15 min at room temperature. Subsequently, the slides were developed with diaminobenzidine and then coun-

Table 1: Correlation of PSCA expression with Gleason score

Gleason score	Intensity × frequency	
	0-6 (%)	9 (%)
2-4	5 (83)	1 (17)
5-7	19 (79)	5 (21)
8-10	5 (28)	13 (72)

Table 2: Correlation of PSCA expression with clinical stage

Tumor stage	Intensity × frequency	
	0-6 (%)	9 (%)
≤B	27 (67.5)	13 (32.5)
≥C	2 (25)	6 (75)

terstained with hematoxylin to localize the hybridization signals. Sections hybridized with the sense control probes routinely did not show any specific hybridization signal above background. All slides were hybridized with PBS to substitute for the probes as a negative control.

#### Scoring methods

To determine the correlation between the results of PSCA immunostaining and mRNA *in situ* hybridization, the same scoring manners are taken in the present study for PSCA protein staining by IHC and PSCA mRNA staining by ISH. Each slide was read and scored by two independently experienced urological pathologists using Olympus BX-41 light microscopes. The evaluation was done in a blinded fashion. For each section, five areas of similar grade were analyzed semiquantitatively for the fraction of cells staining. Fifty percent of specimens were randomly chosen and rescored to determine the degree of interobserver and intraobserver concordance. There was greater than 95% intra- and interobserver agreement.

The intensity of PSCA expression evaluated microscopically was graded on a scale of 0 to 3+ with 3 being the highest expression observed (0, no staining; 1+, mildly intense; 2+, moderately intense; 3+, severely intense). The staining density was quantified as the percentage of cells staining positive for PSCA with the primary antibody or hybridization probe, as follows: 0 = no staining; 1 = positive staining in <25% of the sample; 2 = positive staining in 25%-50% of the sample; 3 = positive staining in >50%

of the sample. Intensity score (0 to 3+) was multiplied by the density score (0-3) to give an overall score of 0-9 [1,5]. In this way, we were able to differentiate specimens that may have had focal areas of increased staining from those that had diffuse areas of increased staining [6]. The overall score for each specimen was then categorically assigned to one of the following groups: 0 score, negative expression; 1-2 scores, weak expression; 3-6 scores, moderate expression; 9 score, strong expression.

#### Statistical analysis

Intensity and density of PSCA protein and mRNA expression in BPH, PIN and Pca tissues were compared using the Chi-square and Student's *t*-test. Univariate associations between PSCA expression and Gleason score, clinical stage and progression to androgen-independence were calculated using Fisher's Exact Test. For all analyses,  $p < 0.05$  was considered statistically significant.

#### Results

##### PSCA expression in BPH

In general, PSCA protein and mRNA were expressed weakly in individual samples of BPH. Some areas of prostate expressed weak levels (composite score 1-2), whereas other areas were completely negative (composite score 0). Four cases (20%) of BPH had moderate expression of PSCA protein and mRNA (composite score 4-6) by IHC and ISH. In 2/20 (10%) BPH specimens, PSCA mRNA expression was moderate (composite score 3-6), but PSCA protein expression was weak (composite score

2) in one and negative (composite score 0) in the other. PSCA expression was localized to the basal and secretory epithelial cells, and prostatic stroma was almost negative staining for PSCA protein and mRNA in all cases examined.

#### **PSCA expression in PIN**

In this study, we detected weak or negative expression of PSCA protein and mRNA ( $\leq 2$  scores) in 7 of 9 (77.8%) low grade PIN and in 2 of 11 (18.2%) HGPIN, and moderate expression (3–6 scores) in the rest 2 low grade PIN and 5 of 11 (45.5%) HGPIN. One HGPIN with moderate PSCA mRNA expression (6 score) was found weak staining for PSCA protein (2 score) by IHC. Strong PSCA protein and mRNA expression (9 score) were detected in the remaining 3 of 11 (27.3%) HGPIN. There was a statistically significant difference of PSCA protein and mRNA expression levels observed between HGPIN and BPH ( $p < 0.05$ ), but no statistical difference reached between low grade PIN and BPH ( $p > 0.05$ ).

#### **PSCA expression in Pca**

In order to determine if PSCA protein and mRNA can be detected in prostate cancers and if PSCA expression levels are increased in malignant compared with benign glands, Forty-eight paraffin-embedded Pca specimens were analysed by IHC and ISH. It was shown that 19 of 48 (39.6%) Pca samples stained very strongly for PSCA protein and mRNA with a score of 9 and another 21 (43.8%) specimens displayed moderate staining with scores of 4–6 (Figure 1). In addition, 4 specimens with moderate to strong PSCA mRNA expression (scores of 4–9) had weak protein staining (a score of 2) by IHC analyses. Overall, Pca expressed a significantly higher level of PSCA protein and mRNA than any other specimen category in this study ( $p < 0.05$ , compared with BPH and PIN respectively). The result demonstrates that PSCA protein and mRNA are overexpressed by a majority of human Pca.

#### **Correlation of PSCA expression with Gleason score in Pca**

Using the semi-quantitative scoring method as described in Materials and Methods, we compared the expression level of PSCA protein and mRNA with Gleason grade of Pca, as shown in Table 1. Prostate adenocarcinomas were graded by Gleason score as 2–4 scores = well-differentiation, 5–7 scores = moderate-differentiation and 8–10 scores = poor-differentiation [7]. Seventy-two percent of Gleason scores 8–10 prostate cancers had very strong staining of PSCA compared to 21% with Gleason scores 5–7 and 17% with 2–4 respectively, demonstrating that poorly differentiated Pca had significantly stronger expression of PSCA protein and mRNA than moderately and well differentiated tumors ( $p < 0.05$ ). As depicted in Figure 1, IHC and ISH analyses showed that PSCA protein and mRNA expression in several cases of poorly differen-

tiated Pca were particularly prominent, with more intense and uniform staining. The results indicate that PSCA expression increases significantly with higher tumor grade in human Pca.

#### **Correlation of PSCA expression with clinical stage in Pca**

With regards to PSCA expression in every stage of Pca, we showed the results in Table 2. Seventy-five percent of locally advanced and node positive cancers (i.e. C-D stages) expressed statistically high levels of PSCA versus 32.5% that were organ confined (i.e. A-B stages) ( $p < 0.05$ ). The data demonstrate that PSCA expression increases significantly with advanced tumor stage in human Pca.

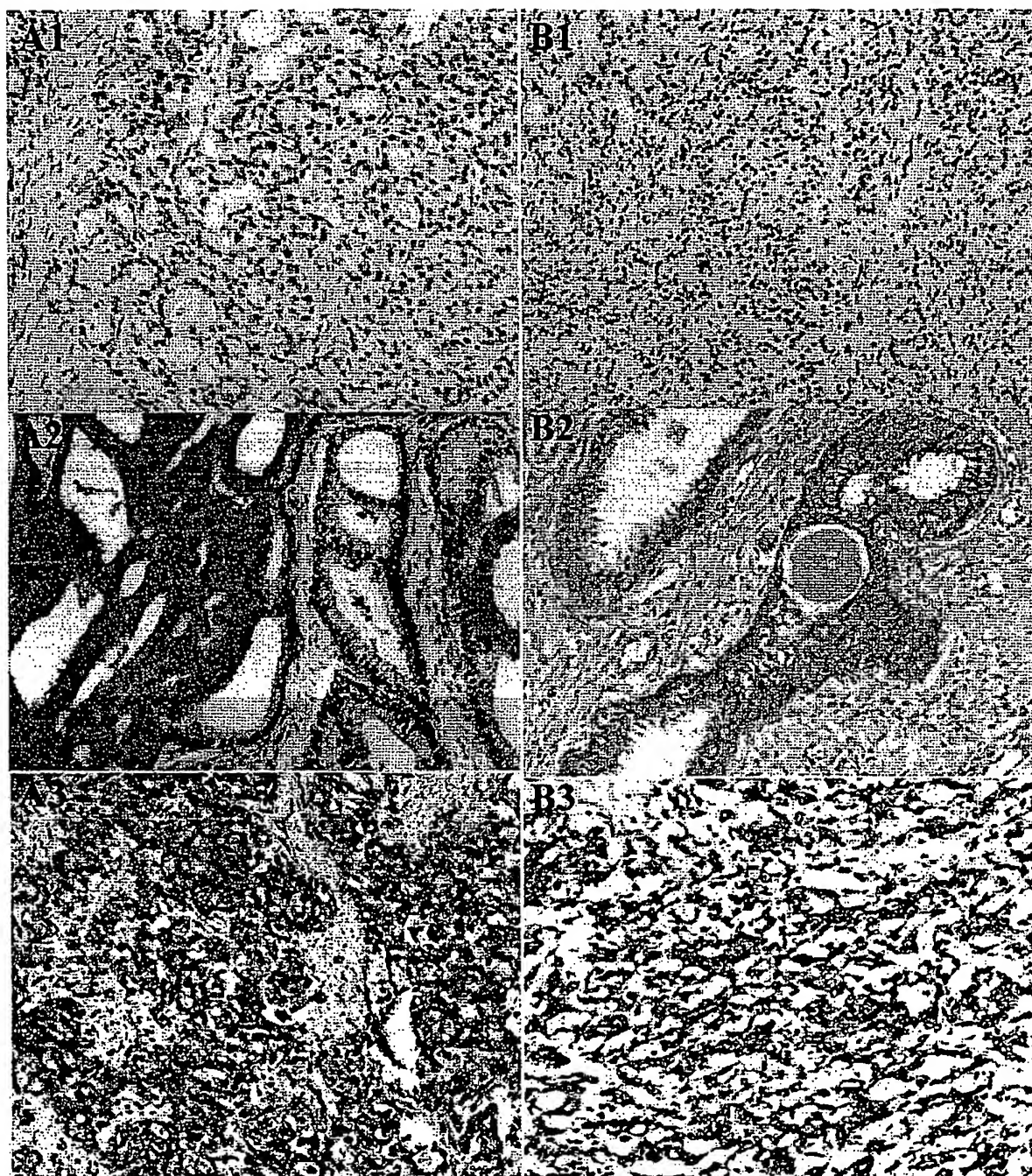
#### **Correlation of PSCA expression with androgen-independent progression of Pca**

All 9 specimens of androgen-independent prostate cancers stained positive for PSCA protein and mRNA. Eight specimens were obtained from patients managed prior to androgen ablation therapy. Seven of eight (87.5%) of these androgen-independent prostate cancers were in the strongest staining category (score = 9), compared with three out of eight (37.5%) of patients with androgen-dependent cancers ( $p < 0.05$ ). The results demonstrate that PSCA expression increases significantly with progression to androgen-independence of human Pca.

It is evident from the results above that within a majority of human prostate cancers the level of PSCA protein and mRNA expression correlates significantly with increasing grade, worsening stage and progression to androgen-independence.

#### **Correlation of PSCA immunostaining and mRNA in situ hybridization**

In all 88 specimens surveyed herein, we compared the results of PSCA IHC staining with mRNA ISH analysis. Positive staining areas and its intensity and density scores evaluated by IHC were identical to those seen by ISH in 79 of 88 (89.8%) specimens (18/20 BPH, 19/20 PIN and 42/48 Pca respectively). Importantly, 27/27 samples with PSCA mRNA composite scores of 0–2, 32/36 samples with scores of 3–6 and 22/24 samples with a score of 9 also had PSCA protein expression scores of 0–2, 3–6 and 9 respectively. However, in 5 samples with PSCA mRNA overall scores of 3–6 and in 2 with scores of 9 there were less or negative PSCA protein expression (i.e. scores of 0–4), suggesting that this may reflect posttranscriptional modification of PSCA or that the epitopes recognized by PSCA mAb may be obscured in some cancers. The data demonstrate that the results of PSCA immunostaining were consistent with those of mRNA ISH analysis, showing a high degree of correlation between PSCA protein and mRNA expression.



**Figure 1**

Representatives of PSCA IHC and ISH staining in Pca (A. IHC staining, B. ISH staining,  $\times 200$  magnification). A<sub>1</sub>, B<sub>1</sub>: negative control of IHC and ISH. PBS replacing the primary antibody (A<sub>1</sub>) and hybridization with a sense PSCA probe (B<sub>1</sub>) showed no background staining. A<sub>2</sub>, B<sub>2</sub>: a moderately differentiated Pca (Gleason score =  $3+3 = 6$ ) with moderate staining (composite score = 6) in all malignant cells; A<sub>2</sub>: IHC shows not only cell surface but also apparent cytoplasmic staining of PSCA protein. A<sub>3</sub>, B<sub>3</sub>: a poorly differentiated Pca (Gleason score =  $4+4 = 8$ ) with very strong staining (composite score = 9) in all malignant cells.

## Discussion

PSCA is homologous to a group of cell surface proteins that mark the earliest phase of hematopoietic development. PSCA mRNA expression is prostate-specific in normal male tissues and is highly up-regulated in both androgen-dependent and-independent Pca xenografts (LAPC-4 tumors). We hypothesize that PSCA may play a role in Pca tumorigenesis and progression, and may serve as a target for Pca diagnosis and treatment. In this study, IHC and ISH showed that in general there were weak or absent PSCA protein and mRNA expression in BPH and low grade PIN tissues. However, PSCA protein and mRNA are widely expressed in HGPIN, the putative precursor of invasive Pca, suggesting that up-regulation of PSCA is an early event in prostate carcinogenesis. Recently, Reiter RE et al [1], using ISH analysis, reported that 97 of 118 (82%) HGPIN specimens stained strongly positive for PSCA mRNA. A very similar finding was seen on mouse PSCA (mPSCA) expression in mouse HGPIN tissues by Tran C. P et al [8]. These data suggest that PSCA may be a new marker associated with transformation of prostate cells and tumorigenesis.

We observed that PSCA protein and mRNA are highly expressed in a large percentage of human prostate cancers, including advanced, poorly differentiated, androgen-independent and metastatic cases. Fluorescence-activated cell sorting and confocal/ immunofluorescent studies demonstrated cell surface expression of PSCA protein in Pca cells [9]. Our IHC expression analysis of PSCA shows not only cell surface but also apparent cytoplasmic staining of PSCA protein in Pca specimens (Figure 1). One possible explanation for this is that anti-PSCA antibody can recognize PSCA peptide precursors that reside in the cytoplasm. Also, it is possible that the positive staining that appears in the cytoplasm is actually from the overlying cell membrane [5]. These data seem to indicate that PSCA is a novel cell surface marker for human Pca.

Our results show that elevated level of PSCA expression correlates with high grade (i.e. poor differentiation), increased tumor stage and progression to androgen-independence of Pca. These findings support the original IHC analyses by Gu Z et al [9], who reported that PSCA protein expressed in 94% of primary Pca and the intensity of PSCA protein expression increased with tumor grade, stage and progression to androgen-independence. Our results also collaborate the recent work of Han KR et al [10], in which the significant association between high PSCA expression and adverse prognostic features such as high Gleason score, seminal vesicle invasion and capsular involvement in Pca was found. It is suggested that PSCA overexpression may be an adverse predictor for recurrence, clinical progression or survival of Pca. Hara H et al [11] used RT-PCR detection of PSA, PSMA and PSCA in 1

ml of peripheral blood to evaluate Pca patients with poor prognosis. The results showed that among 58 Pca patients, each PCR indicated the prognostic value in the hierarchy of PSCA>PSA>PSMA RT-PCR, and extraprostatic cases with positive PSCA PCR indicated lower disease-progression-free survival than those with negative PSCA PCR, demonstrating that PSCA can be used as a prognostic factor. Dubey P et al [12] reported that elevated numbers of PSCA + cells correlate positively with the onset and development of prostate carcinoma over a long time span in the prostates of the TRAMP and PTEN +/- models compared with its normal prostates. Taken together with our present findings, in which PSCA is overexpressed from HGPIN to almost frank carcinoma, it is reasonable and possible to use increased PSCA expression level or increased numbers of PSCA-positive cells in the prostate samples as a prognostic marker to predict the potential onset of this cancer. These data raise the possibility that PSCA may have diagnostic utility or clinical prognostic value in human Pca.

The cause of PSCA overexpression in Pca is not known. One possible mechanism is that it may result from PSCA gene amplification. In humans, PSCA is located on chromosome 8q24.2 [1], which is often amplified in metastatic and recurrent Pca and considered to indicate a poor prognosis [13-15]. Interestingly, PSCA is in close proximity to the c-myc oncogene, which is amplified in >20% of recurrent and metastatic prostate cancers [16,17]. Reiter RE et al [18] reported that PSCA and MYC gene copy numbers were co-amplified in 25% of tumors (five out of twenty), demonstrating that PSCA overexpression is associated with PSCA and MYC coamplification in Pca. Gu Z et al [9] recently reported that in 102 specimens available to compare the results of PSCA immunostaining with their previous mRNA ISH analysis, 92 (90.2%) had identically positive areas of PSCA protein and mRNA expression. Taken together with our findings, in which we detected moderate to strong expression of PSCA protein and mRNA in 34 of 40 (85%) Pca specimens examined simultaneously by IHC and ISH analyses, it is demonstrated that PSCA protein and mRNA overexpressed in human Pca, and that the increased protein level of PSCA was resulted from the upregulated transcription of its mRNA.

At present, the regulation mechanisms of human PSCA expression and its biological function are yet to be elucidated. PSCA expression may be regulated by multiple factors [18]. Watabe T et al [19] reported that transcriptional control is a major component regulating PSCA expression levels. In addition, induction of PSCA expression may be regulated or mediated through cell-cell contact and protein kinase C (PKC) [20]. Homologues of PSCA have diverse activities, and have themselves been involved in

carcinogenesis. Signalling through SCA-2 has been demonstrated to prevent apoptosis in immature thymocytes [21]. Thy-1 is involved in T cell activation and transduces signals through src-like tyrosine kinases [22]. Ly-6 genes have been implicated both in tumorigenesis and in cell-cell adhesion [23-25]. Cell-cell or cell-matrix interaction is critical for local tumor growth and spread to distal sites. From its restricted expression in basal cells of normal prostate and its homology to SCA-2, PSCA may play a role in stem/progenitor cell function, such as self-renewal (i.e. anti-apoptosis) and/or proliferation [1]. Taken together with the results in the present study, we speculate that PSCA may play a role in tumorigenesis and clinical progression of Pca through affecting cell transformation and proliferation. From our results, it is also suggested that PSCA as a new cell surface antigen may have a number of potential uses in the diagnosis, therapy and clinical prognosis of human Pca. PSCA overexpression in prostate biopsies could be used to identify patients at high risk to develop recurrent or metastatic disease, and to discriminate cancers from normal glands in prostatectomy samples. Similarly, the detection of PSCA-overexpressing cells in bone marrow or peripheral blood may identify and predict metastatic progression better than current assays, which identify only PSA-positive or PSMA-positive prostate cells.

In summary, we have shown in this study that PSCA protein and mRNA are maintained in expression from HGPIN through all stages of Pca in a majority of cases, which may be associated with prostate carcinogenesis and correlate positively with high tumor grade (poor cell differentiation), advanced stage and androgen-independent progression. PSCA protein overexpression is due to the upregulation of its mRNA transcription. The results suggest that PSCA may be a promising molecular marker for the clinical prognosis of human Pca and a valuable target for diagnosis and therapy of this tumor.

### Competing interests

None declared.

### References

- Reiter RE, Gu Z, Watabe T, Thomas G, Szigeti K, David E, Wahl M, Nisitani S, Yamashiro J, Le Beau MM, Loda M, Witte ON: Prostate stem cell antigen: a cell surface marker overexpressed in prostate cancer. *Proc Natl Acad Sci USA* 1998, 95:1735-1740.
- Gleason DF: Histologic grading and clinical staging of prostatic carcinoma. In: *Urologic Pathology: The Prostate* Edited by: Tannebaum M. Philadelphia, Lea & Febiger; 1977:171-197.
- Brawer MK: Prostatic intraepithelial neoplasia: a premalignant lesion. *Hum Pathol* 1992, 23:242-248.
- Amin MB, Ro JY, Ayala AC: Prostatic intraepithelial neoplasia: relationship to adenocarcinoma of prostate. *Pathol Annu* 1994, 29:1-30.
- Amara N, Palapattu GS, Schrage M, Gu Z, Thomas GV, Dorey F, Said J, Reiter RE: Prostate stem cell antigen is overexpressed in human transitional cell carcinoma. *Cancer Res* 2001, 61:4660-4665.
- Hanas JS, Lerner MR, Lightfoot SA, Raczkowski C, Kastens DJ, Brackert DJ, Postler RG: Expression of the cyclin-dependent kinase inhibitor p21 (WAF1/CIP1) and p53 tumor suppressor in dysplastic progression and adenocarcinoma in Barrett esophagus. *Cancer (Phila)* 1999, 86:756-763.
- Egevad L, Gramfors T, Karlberg L: Prognostic value of the Gleason score in prostate cancer. *BJU Int* 2002, 89:538-542.
- Tran CP, Lin C, Yamashiro J, Reiter RE: Prostate stem cell antigen is a marker of late intermediate prostate epithelial cells. *Mol Cancer Res* 2002, 1:113-121.
- Gu Z, Thomas G, Yamashiro J, Shintaku IP, Dorey F, Raitano A, Witte ON, Said JW, Loda M, Reiter RE: Prostate stem cell antigen (PSCA) expression increases with high Gleason score, advanced stage and bone metastasis in prostate cancer. *Oncogene* 2000, 19:1288-1296.
- Han KR, Sellgren DB, Liu X, Horvath S, Shintaku PI, Thomas GV, Said JW, Reiter RE: Prostate stem cell antigen expression is associated with gleason score, seminal vesicle invasion and capsular invasion in prostate cancer. *J Urol* 2004, 171:1117-1121.
- Hara H, Kasahara T, Kawasaki T, Bilim V, Obara K, Takahashi K, Tomita Y: Reverse Transcription-Polymerase Chain Reaction Detection of Prostate-specific Antigen, Prostate-specific Membrane Antigen, and Prostate Stem Cell Antigen in One Milliliter of Peripheral Blood. *Clin Cancer Res* 2002, 8:1794-1799.
- Dubey P, Wu H, Reiter RE, Witte ON: Alternative pathways to prostate carcinoma activate prostate stem cell antigen expression. *Cancer Res* 2001, 61:3256-3261.
- Visakorpi T, Kallioniemi AH, Syvanen AC, Hyytinen ER, Karhu R, Tammela T, Isola JJ, Kallioniemi OP: Genetic changes in primary and recurrent prostate cancer by comparative genomic hybridization. *Cancer Res* 1995, 55:342-347.
- Sato K, Qian J, Slezak JM, Lieber MM, Bostwick DG, Bergstrahl EJ, Jenkins RB: Clinical significance of alterations of chromosome 8 in high-grade, advanced, nonmetastatic prostate carcinoma. *J Natl Cancer Inst* 1999, 91:1574-1580.
- Van Den Berg C, Guan XY, Von Hoff D, Jenkins R, Bittner J, Griffin C, Kallioniemi O, Visakorpi T, McGill J, Herath J, Epstein J, Sarosdy M, Meltzer P, Trent J: DNA sequence amplification in human prostate cancer identified by chromosome microdissection: potential prognostic implications. *Clin Cancer Res* 1995, 1:11-18.
- Jenkins RB, Qian J, Lieber MM, Bostwick DG: Detection of c-myc oncogene amplification and chromosomal anomalies in metastatic prostatic carcinoma by fluorescence in situ hybridization. *Cancer Res* 1997, 57:524-531.
- Nupponen NN, Kakkola L, Koivisto P, Visakorpi T: Genetic alterations in hormone-refractory recurrent prostate carcinomas. *Am J Pathol* 1998, 153:141-148.
- Reiter RE, Sato I, Thomas G, Qian J, Gu Z, Watabe T, Loda M, Jenkins RB: Coamplification of prostate stem cell antigen (PSCA) and MYC in locally advanced prostate cancer. *Genes Chromosomes Cancer* 2000, 27:95-103.
- Watabe T, Lin M, Donjacour AA, Cunha GR, Witte ON, Reiter RE: Growth, regeneration, and tumorigenesis of the prostate activates the PSCA promoter. *Proc Natl Acad Sci USA* 2002, 99:401-406.
- Bahrenberg G, Brauers A, Joost HG, Jakse G: PSCA expression is regulated by phorbol ester and cell adhesion in the bladder carcinoma cell line RT112. *Cancer Lett* 2001, 168:37-43.
- Noda S, Kosugi A, Saitoh S, Narumiya S, Hamaoka T: Protection from anti-TCR/CD3-induced apoptosis in immature thymocytes by a signal through thymic shared antigen-1/stem cell antigen-2. *J Exp Med* 1996, 183:2355-2360.
- Thomas PM, Samelson LE: The glycoposphatidylinositol-anchored Thy-1 molecule interacts with the p60fyn protein tyrosine kinase in T cells. *J Biol Chem* 1992, 267:12317-12322.
- Bamezai A, Rock KL: Overexpressed Ly-6A.2 mediated cell-cell adhesion by binding a ligand expressed on lymphoid cells. *Proc Natl Acad Sci USA* 1995, 92:4294-4298.
- Katz BZ, Eshel R, Sagi-Assif O, Witz IP: An association between high Ly-6A/E expression on tumor cells and a highly malignant phenotype. *Int J Cancer* 1994, 59:684-691.
- Brakenhoff RH, Gerretsen M, Knippels EM, van Dijk M, van Essen H, Weghuis DO, Sinke RJ, Snow GB, van Dongen GA: The human E48 antigen, highly homologous to the murine Ly-6 antigen ThB, is a GPI-anchored molecule apparently involved in keratinocyte cell-cell adhesion. *J Cell Biol* 1995, 129:1677-1689.



## Review

# Translation Initiation in Cancer: A Novel Target for Therapy<sup>1</sup>

Funda Meric<sup>2</sup> and Kelly K. Hunt

Department of Surgical Oncology, The University of Texas M. D. Anderson Cancer Center, Houston, Texas 77030

### Abstract

Translation initiation is regulated in response to nutrient availability and mitogenic stimulation and is coupled with cell cycle progression and cell growth. Several alterations in translational control occur in cancer. Variant mRNA sequences can alter the translational efficiency of individual mRNA molecules, which in turn play a role in cancer biology. Changes in the expression or availability of components of the translational machinery and in the activation of translation through signal transduction pathways can lead to more global changes, such as an increase in the overall rate of protein synthesis and translational activation of the mRNA molecules involved in cell growth and proliferation. We review the basic principles of translational control, the alterations encountered in cancer, and selected therapies targeting translation initiation to help elucidate new therapeutic avenues.

### Introduction

The fundamental principle of molecular therapeutics in cancer is to exploit the differences in gene expression between cancer cells and normal cells. With the advent of cDNA array technology, most efforts have concentrated on identifying differences in gene expression at the level of mRNA, which can be attributable either to DNA amplification or to differences in transcription. Gene expression is quite complicated, however, and is also regulated at the level of mRNA stability, mRNA translation, and protein stability.

The power of translational regulation has been best recognized among developmental biologists, because transcription does not occur in early embryogenesis in eukaryotes. For example, in *Xenopus*, the period of transcriptional quiescence continues until the embryo reaches midblastula transition, the 4000-cell stage. Therefore, all necessary mRNA molecules are transcribed during oogenesis and stockpiled in a translationally inactive, masked form. The mRNA are translationally activated at appropriate times during oocyte maturation, fertilization, and

early embryogenesis and thus, are under strict translational control.

Translation has an established role in cell growth. Basically, an increase in protein synthesis occurs as a consequence of mitogenesis. Until recently, however, little was known about the alterations in mRNA translation in cancer, and much is yet to be discovered about their role in the development and progression of cancer. Here we review the basic principles of translational control, the alterations encountered in cancer, and selected therapies targeting translation initiation to elucidate potential new therapeutic avenues.

### Basic Principles of Translational Control

#### Mechanism of Translation Initiation

Translation initiation is the main step in translational regulation. Translation initiation is a complex process in which the initiator tRNA and the 40S and 60S ribosomal subunits are recruited to the 5' end of a mRNA molecule and assembled by eukaryotic translation initiation factors into an 80S ribosome at the start codon of the mRNA (Fig. 1). The 5' end of eukaryotic mRNA is capped, i.e., contains the cap structure m<sup>7</sup>GpppN (7-methyl-guanosine-triphospho-5'-ribonucleoside). Most translation in eukaryotes occurs in a cap-dependent fashion, i.e., the cap is specifically recognized by the eIF4E,<sup>3</sup> which binds the 5' cap. The eIF4F translation initiation complex is then formed by the assembly of eIF4E, the RNA helicase eIF4A, and eIF4G, a scaffolding protein that mediates the binding of the 40S ribosomal subunit to the mRNA molecule through interaction with the eIF3 protein present on the 40S ribosome. eIF4A and eIF4B participate in melting the secondary structure of the 5' UTR of the mRNA. The 43S initiation complex (40S/eIF2/Met-tRNA/GTP complex) scans the mRNA in a 5'→3' direction until it encounters an AUG start codon. This start codon is then base-paired to the anticodon of initiator tRNA, forming the 48S initiation complex. The initiation factors are then displaced from the 48S complex, and the 60S ribosome joins to form the 80S ribosome.

Unlike most eukaryotic translation, translation initiation of certain mRNAs, such as the picornavirus RNA, is cap independent and occurs by internal ribosome entry. This mechanism does not require eIF4E. Either the 43S complex can bind the initiation codon directly through interaction with the IRES in the 5' UTR such as in the encephalomyocarditis virus, or it can

Received 5/16/02; revised 7/12/02; accepted 7/22/02.

<sup>1</sup> F. M. is supported by The University of Texas M. D. Anderson Cancer Center Physician-Scientist Program and by NIH Grant 1K08-CA 91895-01. K. K. H. is supported by Department of Defense Award DAMD-17-97-1-7162.

<sup>2</sup> To whom requests for reprints should be addressed, at Department of Surgical Oncology, Box 444, The University of Texas M. D. Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, TX 77030. Phone: (713) 745-4453; Fax: (713) 745-4926; E-mail: fmeric@mdanderson.org.

<sup>3</sup> The abbreviations used are: eIF4E, eukaryotic initiation factor 4E; UTR, untranslated region; IRES, internal ribosome entry site; 4E-BP1, eukaryotic initiation factor 4E-binding protein 1; S6K, ribosomal p70 S6 kinase; mTOR, mammalian target of rapamycin; ATM, ataxia telangiectasia mutated; PI3K, phosphatidylinositol 3-kinase; PTEN, phosphatase and tensin homolog deleted from chromosome 10; PP2A, protein phosphatase 2A; TGF- $\beta$ 3, transforming growth factor- $\beta$ 3; PAP, poly(A) polymerase; EPA, eicosapentaenoic acid; mda-7, melanoma differentiation-associated gene 7.

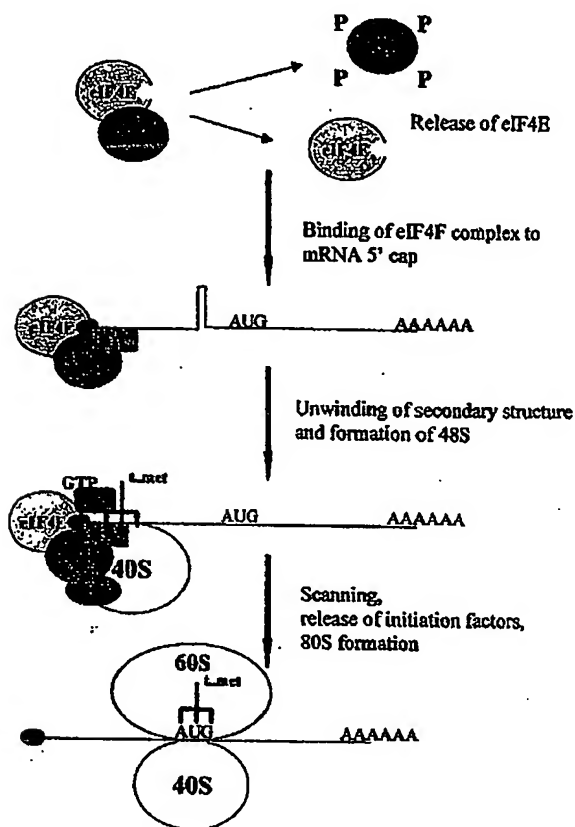


Fig. 1. Translation initiation in eukaryotes. The 4E-BPs are hyperphosphorylated to release eIF4E so that it can interact with the 5' cap, and the eIF4F initiation complex is assembled. The interaction of poly(A) binding protein with the initiation complex and circularization of the mRNA is not depicted in the diagram. The secondary structure of the 5' UTR is melted, the 40S ribosomal subunit is bound to eIF3, and the ternary complex consisting of eIF2, GTP, and the Met-tRNA are recruited to the mRNA. The ribosome scans the mRNA in a 5'→3' direction until an AUG start codon is found in the appropriate sequence context. The initiation factors are released, and the large ribosomal subunit is recruited.

Initially attach to the IRES and then reach the initiation codon by scanning or transfer, as is the case with the poliovirus (1).

### Regulation of Translation Initiation

Translation initiation can be regulated by alterations in the expression or phosphorylation status of the various factors involved. Key components in translational regulation that may provide potential therapeutic targets follow.

**eIF4E.** eIF4E plays a central role in translation regulation. It is the least abundant of the initiation factors and is considered the rate-limiting component for initiation of cap-dependent translation. eIF4E may also be involved in mRNA splicing, mRNA 3' processing, and mRNA nucleocytoplasmic transport (2). eIF4E expression can be increased at the transcriptional level in response to serum or growth factors (3). eIF4E overexpression may cause preferential translation of mRNAs containing excessive secondary structure in their 5' UTR that are normally discriminated against by the trans-

lational machinery and thus are inefficiently translated (4–7). As examples of this, overexpression of eIF4E promotes increased translation of vascular endothelial growth factor, fibroblast growth factor-2, and cyclin D1 (2, 8, 9).

Another mechanism of control is the regulation of eIF4E phosphorylation. eIF4E phosphorylation is mediated by the mitogen-activated protein kinase-interacting kinase 1, which is activated by the mitogen-activated pathway activating extracellular signal-related kinases and the stress-activated pathway acting through p38 mitogen-activated protein kinase (10–13). Several mitogens, such as serum, platelet-derived growth factor, epidermal growth factor, insulin, angiotensin II, src kinase overexpression, and ras overexpression, lead to eIF4E phosphorylation (14). The phosphorylation status of eIF4E is usually correlated with the translational rate and growth status of the cell; however, eIF4E phosphorylation has also been observed in response to some cellular stresses when translational rates actually decrease (15). Thus, further study is needed to understand the effects of eIF4E phosphorylation on eIF4E activity.

Another mechanism of regulation is the alteration of eIF4E availability by the binding of eIF4E to the eIF4E-binding proteins (4E-BP, also known as PHAS-I). 4E-BPs compete with eIF4G for a binding site in eIF4E. The binding of eIF4E to the best characterized eIF4E-binding protein, 4E-BP1, is regulated by 4E-BP1 phosphorylation. Hypophosphorylated 4E-BP1 binds to eIF4E, whereas 4E-BP1 hyperphosphorylation decreases this binding. Insulin, angiotensin, epidermal growth factor, platelet-derived growth factor, hepatocyte growth factor, nerve growth factor, insulin-like growth factors I and II, interleukin 3, granulocyte-macrophage colony-stimulating factor + steel factor, gastrin, and the adenovirus have all been reported to induce phosphorylation of 4E-BP1 and to decrease the ability of 4E-BP1 to bind eIF4E (15, 16). Conversely, deprivation of nutrients or growth factors results in 4E-BP1 dephosphorylation, an increase in eIF4E binding, and a decrease in cap-dependent translation.

**p70 S6 Kinase.** Phosphorylation of ribosomal 40S protein S6 by S6K is thought to play an important role in translational regulation. S6K  $-/-$  mouse embryonic cells proliferate more slowly than do parental cells, demonstrating that S6K has a positive influence on cell proliferation (17). S6K regulates the translation of a group of mRNAs possessing a 5' terminal oligopyrimidine tract (5' TOP) found at the 5' UTR of ribosomal protein mRNAs and other mRNAs coding for components of the translational machinery. Phosphorylation of S6K is regulated in part based on the availability of nutrients (18, 19) and is stimulated by several growth factors, such as platelet-derived growth factor and insulin-like growth factor I (20).

**eIF2 $\alpha$  Phosphorylation.** The binding of the initiator tRNA to the small ribosomal unit is mediated by translation initiation factor eIF2. Phosphorylation of the  $\alpha$ -subunit of eIF2 prevents formation of the eIF2/GTP/Met-tRNA complex and inhibits global protein synthesis (21, 22). eIF2 $\alpha$  is phosphorylated under a variety of conditions, such as viral infection, nutrient deprivation, heme deprivation, and apoptosis (22). eIF2 $\alpha$  is phosphorylated by heme-regulated inhibitor, nutrient-regulated protein kinase, and the IFN-induced, double-stranded RNA-activated protein kinase (PKR; Ref. 23).



**The mTOR Signaling Pathway.** The macrolide antibiotic rapamycin (Siralimus; Wyeth-Ayerst Research, Collegeville, PA) has been the subject of intensive study because it inhibits signal transduction pathways involved in T-cell activation. The rapamycin-sensitive component of these pathways is mTOR (also called FRAP or RAFT1). mTOR is the mammalian homologue of the yeast TOR proteins that regulate  $G_1$  progression and translation in response to nutrient availability (24). mTOR is a serine-threonine kinase that modulates translation initiation by altering the phosphorylation status of 4E-BP1 and S6K (Fig. 2; Ref. 25).

4E-BP1 is phosphorylated on multiple residues. mTOR phosphorylates the Thr-37 and Thr-46 residues of 4E-BP1 *in vitro* (26); however, phosphorylation at these sites is not associated with a loss of eIF4E binding. Phosphorylation of Thr-37 and Thr-46 is required for subsequent phosphorylation at several COOH-terminal, serum-sensitive sites; a combination of these phosphorylation events appears to be needed to inhibit the binding of 4E-BP1 to eIF4E (25). The product of the *ATM* gene, p38/MSK1 pathway, and protein kinase *C* $\alpha$  also play a role in 4E-BP1 phosphorylation (27–29).

S6K and 4E-BP1 are also regulated, in part, by PI3K and its downstream protein kinase Akt. PTEN is a phosphatase that negatively regulates PI3K signaling. PTEN null cells have constitutively active Akt, with increased S6K activity and S6 phosphorylation (30). S6K activity is inhibited both by PI3K inhibitors wortmannin and LY294002 and by mTOR inhibitor rapamycin (24). Akt phosphorylates Ser-2448 in mTOR *in vitro*, and this site is phosphorylated upon Akt activation *in vivo* (31–33). Thus, mTOR is regulated by the PI3K/Akt pathway; however, this does not appear to be the only mode of regulation of mTOR activity. Whether the PI3K pathway also regulates S6K and 4E-BP1 phosphorylation independent of mTOR is controversial.

Interestingly, mTOR autophosphorylation is blocked by wortmannin but not by rapamycin (34). This seeming inconsistency suggests that mTOR-responsive regulation of 4E-BP1 and S6K activity occurs through a mechanism other than intrinsic mTOR kinase activity. An alternate pathway for 4E-BP1 and S6K phosphorylation by mTOR activity is by the inhibition of a phosphatase. Treatment with calyculin A, an inhibitor of phosphatases 1 and 2A, reduces rapamycin-induced dephosphorylation of 4E-BP1 and S6K by rapamycin (35). PP2A interacts with full-length S6K but not with a S6K mutant that is resistant to dephosphorylation resulting from rapamycin. mTOR phosphorylates PP2A *in vitro*; however, how this process alters PP2A activity is not known. These results are consistent with the model that phosphorylation of a phosphatase by mTOR prevents dephosphorylation of 4E-BP1 and S6K, and conversely, that nutrient deprivation and rapamycin block inhibition of the phosphatase by mTOR.

**Polyadenylation.** The poly(A) tail in eukaryotic mRNA is important in enhancing translation initiation and mRNA stability. Polyadenylation plays a key role in regulating gene expression during oogenesis and early embryogenesis. Some mRNA that are translationally inactive in the oocyte are polyadenylated concomitantly with translational activation in oocyte maturation, whereas other mRNAs that are translationally active during oogenesis are deadenylated and trans-

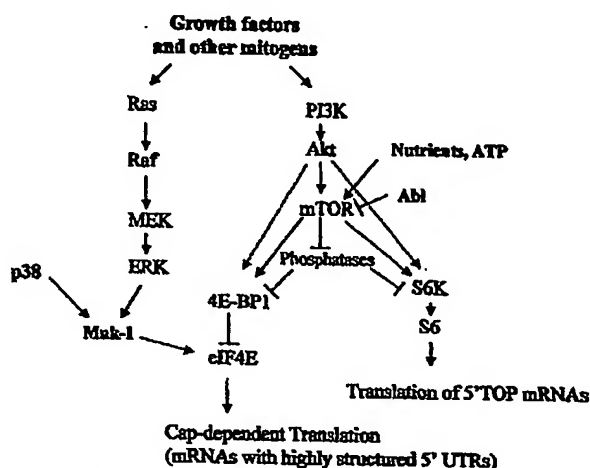


Fig. 2. Regulation of translation initiation by signal transduction pathways. Signaling via p38, extracellular signal-related kinase, PI3K, and mTOR can all activate translation initiation.

lationally silenced (36–38). Thus, control of poly(A) tail synthesis is an important regulatory step in gene expression. The 5' cap and poly(A) tail are thought to function synergistically to regulate mRNA translational efficiency (39, 40).

**RNA Packaging.** Most RNA-binding proteins are assembled on a transcript at the time of transcription, thus determining the translational fate of the transcript (41). A highly conserved family of Y-box proteins is found in cytoplasmic messenger ribonucleoprotein particles, where the proteins are thought to play a role in restricting the recruitment of mRNA to the translational machinery (41–43). The major mRNA-associated protein, YB-1, destabilizes the interaction of eIF4E and the 5' mRNA cap *in vitro*, and overexpression of YB-1 results in translational repression *in vivo* (44). Thus, alterations in RNA packaging can also play an important role in translational regulation.

### Translation Alterations Encountered in Cancer

Three main alterations at the translational level occur in cancer: variations in mRNA sequences that increase or decrease translational efficiency, changes in the expression or availability of components of the translational machinery, and activation of translation through aberrantly activated signal transduction pathways. The first alteration affects the translation of an individual mRNA that may play a role in carcinogenesis. The second and third alterations can lead to more global changes, such as an increase in the overall rate of protein synthesis, and the translational activation of several mRNA species.

### Variations in mRNA Sequence

Variations in mRNA sequence affect the translational efficiency of the transcript. A brief description of these variations and examples of each mechanism follow.

**Mutations.** Mutations in the mRNA sequence, especially in the 5' UTR, can alter its translational efficiency, as seen in the following examples.

**c-myc.** Saito *et al.* proposed that translation of full-length c-myc is repressed, whereas in several Burkitt lymphomas that have deletions of the mRNA 5' UTR, translation of c-myc is more efficient (45). More recently, it was reported that the 5' UTR of c-myc contains an IRES, and thus c-myc translation can be initiated by a cap-independent as well as a cap-dependent mechanism (46, 47). In patients with multiple myeloma, a C→T mutation in the c-myc IRES was identified (48) and found to cause an enhanced initiation of translation via internal ribosomal entry (49).

**BRCA1.** A somatic point mutation (117 G→C) in position -3 with respect to the start codon of the BRCA1 gene was identified in a highly aggressive sporadic breast cancer (50). Chimeric constructs consisting of the wild-type or mutated BRCA1 5' UTR and a downstream luciferase reporter demonstrated a decrease in the translational efficiency with the 5' UTR mutation.

**Cyclin-dependent Kinase Inhibitor 2A.** Some inherited melanoma kindreds have a G→T transversion at base -34 of cyclin-dependent kinase inhibitor-2A, which encodes a cyclin-dependent kinase 4/cyclin-dependent kinase 6 kinase inhibitor important in G<sub>1</sub> checkpoint regulation (51). This mutation gives rise to a novel AUG translation initiation codon, creating an upstream open reading frame that competes for scanning ribosomes and decreases translation from the wild-type AUG.

**Alternate Splicing and Alternate Transcription Start Sites.** Alterations in splicing and alternate transcription sites can lead to variations in 5' UTR sequence, length, and secondary structure, ultimately impacting translational efficiency.

**ATM.** The ATM gene has four noncoding exons in its 5' UTR that undergo extensive alternative splicing (52). The contents of 12 different 5' UTRs that show considerable diversity in length and sequence have been identified. These divergent 5' leader sequences play an important role in the translational regulation of the ATM gene.

**mdm.** In a subset of tumors, overexpression of the oncoprotein mdm2 results in enhanced translation of the mdm2 mRNA. Use of different promoters leads to two mdm2 transcripts that differ only in their 5' leaders (53). The longer 5' UTR contains two upstream open reading frames, and this mRNA is loaded with ribosomes inefficiently compared with the short 5' UTR.

**BRCA1.** In a normal mammary gland, BRCA1 mRNA is expressed with a shorter leader sequence (5'UTRa), whereas in sporadic breast cancer tissue, BRCA1 mRNA is expressed with a longer leader sequence (5' UTRb); the translational efficiency of transcripts containing 5' UTRb is 10 times lower than that of transcripts containing 5' UTRa (54).

**TGF-β3.** TGF-β3 mRNA includes a 1.1-kb 5' UTR, which exerts an inhibitory effect on translation. Many human breast cancer cell lines contain a novel TGF-β3 transcript with a 5' UTR that is 870 nucleotides shorter and has a 7-fold greater translational efficiency than the normal TGF-β3 mRNA (55).

**Alternate Polyadenylation Sites.** Multiple polyadenylation signals leading to the generation of several transcripts with differing 3' UTR have been described for several mRNA species, such as the RET proto-oncogene (56), ATM gene (52), tissue inhibitor of metalloproteinases-3 (57), RHOA

proto-oncogene (58), and calmodulin-1 (59). Although the effect of these alternate 3' UTRs on translation is not yet known, they may be important in RNA-protein interactions that affect translational recruitment. The role of these alterations in cancer development and progression is unknown.

### Alterations in the Components of the Translation Machinery

Alterations in the components of translation machinery can take many forms.

**Overexpression of eIF4E.** Overexpression of eIF4E causes malignant transformation in rodent cells (60) and the deregulation of HeLa cell growth (61). Polunovsky *et al.* (62) found that eIF4E overexpression substitutes for serum and individual growth factors in preserving viability of fibroblasts, which suggests that eIF4E can mediate both proliferative and survival signaling.

Elevated levels of eIF4E mRNA have been found in a broad spectrum of transformed cell lines (63). eIF4E levels are elevated in all ductal carcinoma *in situ* specimens and invasive ductal carcinomas, compared with benign breast specimens evaluated with Western blot analysis (64, 65). Preliminary studies suggest that this overexpression is attributable to gene amplification (66).

There are accumulating data suggesting that eIF4E overexpression can be valuable as a prognostic marker. eIF4E overexpression was found in a retrospective study to be a marker of poor prognosis in stages I to III breast carcinoma (67). Verification of the prognostic value of eIF4E in breast cancer is now under way in a prospective trial (67). However, in a different study, eIF4E expression was correlated with the aggressive behavior of non-Hodgkin's lymphomas (68). In a prospective analysis of patients with head and neck cancer, elevated levels of eIF4E in histologically tumor-free surgical margins predicted a significantly increased risk of local-regional recurrence (9). These results all suggest that eIF4E overexpression can be used to select patients who might benefit from more aggressive systemic therapy. Furthermore, the head and neck cancer data suggest that eIF4E overexpression is a field defect and can be used to guide local therapy.

**Alterations in Other Initiation Factors.** Alterations in a number of other initiation factors have been associated with cancer. Overproduction of eIF4G, similar to eIF4E, leads to malignant transformation *in vitro* (69). eIF-2α is found in increased levels in bronchioloalveolar carcinomas of the lung (3). Initiation factor eIF-4A1 is overexpressed in melanoma (70) and hepatocellular carcinoma (71). The p40 subunit of translation initiation factor 3 is amplified and overexpressed in breast and prostate cancer (72), and the eIF3-p110 subunit is overexpressed in testicular seminoma (73). The role that overexpression of these initiation factors plays on the development and progression of cancer, if any, is not known.

**Overexpression of S6K.** S6K is amplified and highly overexpressed in the MCF7 breast cancer cell line, compared with normal mammary epithelium (74). In a study by Barlund *et al.* (74), S6K was amplified in 59 of 668 primary breast tumors, and a statistically significant association was observed between amplification and poor prognosis.

**Overexpression of PAP.** PAP catalyzes 3' poly(A) synthesis. PAP is overexpressed in human cancer cells compared with normal and virally transformed cells (75). PAP enzymatic activity in breast tumors has been correlated with PAP protein levels (76) and, in mammary tumor cytosols, was found to be an independent factor for predicting survival (76). Little is known, however, about how PAP expression or activity affects the translational profile.

**Alterations in RNA-binding Proteins.** Even less is known about alterations in RNA packaging in cancer. Increased expression and nuclear localization of the RNA-binding protein YB-1 are indicators of a poor prognosis for breast cancer (77), non-small cell lung cancer (78), and ovarian cancer (79). However, this effect may be mediated at least in part at the level of transcription, because YB-1 increases chemoresistance by enhancing the transcription of a multidrug resistance gene (80).

#### Activation of Signal Transduction Pathways

Activation of signal transduction pathways by loss of tumor suppressor genes or overexpression of certain tyrosine kinases can contribute to the growth and aggressiveness of tumors. An important mutant in human cancers is the tumor suppressor gene *PTEN*, which leads to the activation of the PI3K/Akt pathway. Activation of PI3K and Akt induces the oncogenic transformation of chicken embryo fibroblasts. The transformed cells show constitutive phosphorylation of S6K and of 4E-BP1 (81). A mutant Akt that retains kinase activity but does not phosphorylate S6K or 4E-BP1 does not transform fibroblasts, which suggests a correlation between the oncogenicity of PI3K and Akt and the phosphorylation of S6K and 4E-BP1 (81).

Several tyrosine kinases such as platelet-derived growth factor, insulin-like growth factor, HER2/neu, and epidermal growth factor receptor are overexpressed in cancer. Because these kinases activate downstream signal transduction pathways known to alter translation initiation, activation of translation is likely to contribute to the growth and aggressiveness of these tumors. Furthermore, the mRNA for many of these kinases themselves are under translational control. For example, HER2/neu mRNA is translationally controlled both by a short upstream open reading frame that represses HER2/neu translation in a cell type-independent manner and by a distinct cell type-dependent mechanism that increases translational efficiency (82). HER2/neu translation is different in transformed and normal cells. Thus, it is possible that alterations at the translational level can in part account for the discrepancy between *HER2/neu* gene amplification detected by fluorescence *in situ* hybridization and protein levels detected by immunohistochemical assays.

#### Translation Targets of Selected Cancer Therapy

Components of the translation machinery and signal pathways involved in the activation of translation initiation represent good targets for cancer therapy.

##### Targeting the mTOR Signaling Pathway: Rapamycin and Temsirolumab

Rapamycin inhibits the proliferation of lymphocytes. It was initially developed as an immunosuppressive drug for organ

transplantation. Rapamycin with FKBP 12 (FK506-binding protein,  $M_r$  12,000) binds to mTOR to inhibit its function.

Rapamycin causes a small but significant reduction in the initiation rate of protein synthesis (83). It blocks cell growth in part by blocking S6 phosphorylation and selectively suppressing the translation of 5' TOP mRNAs, such as ribosomal proteins, and elongation factors (83–85). Rapamycin also blocks 4E-BP1 phosphorylation and inhibits cap-dependent but not cap-independent translation (17, 86).

The rapamycin-sensitive signal transduction pathway, activated during malignant transformation and cancer progression, is now being studied as a target for cancer therapy (87). Prostate, breast, small cell lung, glioblastoma, melanoma, and T-cell leukemia are among the cancer lines most sensitive to the rapamycin analogue CCI-779 (Wyeth-Ayerst Research; Ref. 87). In rhabdomyosarcoma cell lines, rapamycin is either cytostatic or cytotoxic, depending on the p53 status of the cell; p53 wild-type cells treated with rapamycin arrest in the G<sub>1</sub> phase and maintain their viability, whereas p53 mutant cells accumulate in G<sub>1</sub> and undergo apoptosis (88, 89). In a recently reported study using human primitive neuroectodermal tumor and medulloblastoma models, rapamycin exhibited more cytotoxicity in combination with cisplatin and camptothecin than as a single agent. *In vivo*, CCI-779 delayed growth of xenografts by 160% after 1 week of therapy and 240% after 2 weeks. A single high-dose administration caused a 37% decrease in tumor volume. Growth inhibition *in vivo* was 1.3 times greater, with cisplatin in combination with CCI-779 than with cisplatin alone (90). Thus, preclinical studies suggest that rapamycin analogues are useful as single agents and in combination with chemotherapy.

Rapamycin analogues CCI-779 and RAD001 (Novartis, Basel, Switzerland) are now in clinical trials. Because of the known effect of rapamycin on lymphocyte proliferation, a potential problem with rapamycin analogues is immunosuppression. However, although prolonged immunosuppression can result from rapamycin and CCI-779 administered on continuous-dose schedules, the immunosuppressive effects of rapamycin analogues resolve in ~24 h after therapy (91). The principal toxicities of CCI-779 have included dermatological toxicity, myelosuppression, infection, mucositis, diarrhea, reversible elevations in liver function tests, hyperglycemia, hypokalemia, hypocalcemia, and depression (87, 92–94). Phase II trials of CCI-779 have been conducted in advanced renal cell carcinoma and in stage II/IV breast carcinoma patients who failed with prior chemotherapy. In the results reported in abstract form, although there were no complete responses, partial responses were documented in both renal cell carcinoma and in breast carcinoma (94, 95). Thus, CCI-779 has documented preliminary clinical activity in a previously treated, unselected patient population.

Active investigation is under way into patient selection for mTOR inhibitors. Several studies have found an enhanced efficacy of CCI-779 in PTEN-null tumors (30, 96). Another study found that six of eight breast cancer cell lines were responsive to CCI-779, although only two of these lines lacked PTEN (97). There was, however, a positive correlation between Akt activation and CCI-779 sensitivity (97). This correlation suggests that activation of the PI3K-Akt pathway,

regardless of whether it is attributable to a PTEN mutation or to overexpression of receptor tyrosine kinases, makes cancer cell amenable to mTOR-directed therapy. In contrast, lower levels of the target of mTOR, 4E-BP1, are associated with rapamycin resistance; thus, a lower 4E-BP1/elf4E ratio may predict rapamycin resistance (98).

Another mode of activity for rapamycin and its analogues appears to be through inhibition of angiogenesis. This activity may be both through direct inhibition of endothelial cell proliferation as a result of mTOR inhibition in these cells or by inhibition of translation of such proangiogenic factors as vascular endothelial growth factor in tumor cells (99, 100).

The angiogenesis inhibitor turostatin, another anticancer drug currently under study, was also found recently to inhibit translation in endothelial cells (101). Through a requisite interaction with Integrin, turostatin inhibits activation of the PI3K/Akt pathway and mTOR in endothelial cells and prevents dissociation of elf4E from 4E-BP1, thereby inhibiting cap-dependent translation. These findings suggest that endothelial cells are especially sensitive to therapies targeting the mTOR-signaling pathway.

#### Targeting elf2 $\alpha$ EPA, Clotrimazole, mda-7, and Flavonoids

EPA is an n-3 polyunsaturated fatty acid found in the fish-based diets of populations having a low incidence of cancer (102). EPA inhibits the proliferation of cancer cells (103), as well as in animal models (104, 105). It blocks cell division by inhibiting translation initiation (105). EPA releases Ca<sup>2+</sup> from intracellular stores while inhibiting their refilling, thereby activating PKR. PKR, in turn phosphorylates and inhibits elf2 $\alpha$ , resulting in the inhibition of protein synthesis at the level of translation initiation. Similarly, clotrimazole, a potent antiproliferative agent *in vitro* and *in vivo*, inhibits cell growth through depletion of Ca<sup>2+</sup> stores, activation of PKR, and phosphorylation of elf2 $\alpha$  (106). Consequently, clotrimazole preferentially decreases the expression of cyclins A, E, and D1, resulting in blockage of the cell cycle in G<sub>1</sub>.

mda-7 is a novel tumor suppressor gene being developed as a gene therapy agent. Adenoviral transfer of mda-7 (Ad-mda7) induces apoptosis in many cancer cells including breast, colorectal, and lung cancer (107–109). Ad-mda7 also induces and activates PKR, which leads to phosphorylation of elf2 $\alpha$  and induction of apoptosis (110).

Flavonoids such as genistein and quercetin suppress tumor cell growth. All three mammalian elf2 $\alpha$  kinases, PKR, heme-regulated inhibitor, and PERK/PEK, are activated by flavonoids, with phosphorylation of elf2 $\alpha$  and inhibition of protein synthesis (111).

#### Targeting elf4A and elf4E: Antisense RNA and Peptides

Antisense expression of elf4A decreases the proliferation rate of melanoma cells (112). Sequestration of elf4E by overexpression of 4E-BP1 is proapoptotic and decreases tumorigenicity (113, 114). Reduction of elf4E with antisense RNA decreases soft agar growth, increases tumor latency, and increases the rates of tumor doubling times (7). Antisense elf4E treat-

ment also reduces the expression of angiogenic factors (115) and has been proposed as a potential adjuvant therapy for head and neck cancers, particularly when elevated elf4E is found in surgical margins. Small molecule inhibitors that bind the elf4G/4E-BP1-binding domain of elf4E are proapoptotic (116) and are also being actively pursued.

#### Exploiting Selective Translation for Gene Therapy

A different therapeutic approach that takes advantage of the enhanced cap-dependent translation in cancer cells is the use of gene therapy vectors encoding suicide genes with highly structured 5' UTR. These mRNA would thus be at a competitive disadvantage in normal cells and not translate well, whereas in cancer cells, they would translate more efficiently. For example, the introduction of the 5' UTR of fibroblast growth factor-2 5' to the coding sequence of *herpes simplex virus type-1 thymidine kinase* gene, allows for selective translation of *herpes simplex virus type-1 thymidine kinase* gene in breast cancer cell lines compared with normal mammary cell lines and results in selective sensitivity to ganciclovir (117).

#### Toward the Future

Translation is a crucial process in every cell. However, several alterations in translational control occur in cancer. Cancer cells appear to need an aberrantly activated translational state for survival, thus allowing the targeting of translation initiation with surprisingly low toxicity. Components of the translational machinery, such as elf4E, and signal transduction pathways involved in translation initiation, such mTOR, represent promising targets for cancer therapy. Inhibitors of the mTOR have already shown some preliminary activity in clinical trials. It is possible that with the development of better predictive markers and better patient selection, response rates to single-agent therapy can be improved. Similar to other cytostatic agents, however, mTOR inhibitors are most likely to achieve clinical utility in combination therapy. In the interim, our increasing understanding of translation initiation and signal transduction pathways promise to lead to the identification of new therapeutic targets in the near future.

#### Acknowledgments

We thank Gayle Nesom from The University of Texas M. D. Anderson Cancer Center Department of Scientific Publications for editorial assistance and Dr. Elmer Bernstein for assistance with manuscript preparation.

#### References

1. Pestova, T. V., Kolupasova, V. G., Lomakin, I. B., Papisenko, E. V., Shatsky, I. N., Agol, V. I., and Hellen, C. U. Molecular mechanisms of translation initiation in eukaryotes. *Proc. Natl. Acad. Sci. USA*, 98: 7028–7036, 2001.
2. Rosenwald, I. B., Kaspar, R., Rousseau, D., Gehrke, L., LeBoulch, P., Chen, J. J., Schmidt, E. V., Sonenberg, N., and London, I. M. Eukaryotic translation initiation factor 4E regulates expression of cyclin D1 at transcriptional and post-transcriptional levels. *J. Biol. Chem.*, 270: 21176–21180, 1995.
3. Rosenwald, I. B., Hutzler, M. J., Wang, S., Savas, L., and Fraire, A. E. Expression of eukaryotic translation initiation factors 4E and 2 $\alpha$  is increased frequently in bronchioloalveolar but not in squamous cell carcinomas of the lung. *Cancer (Phila.)*, 92: 2164–2171, 2001.

4. Darveau, A., Pelletier, J., and Sonenberg, N. Differential efficiencies of *in vitro* translation of mouse c-myc transcripts differing in the 5' untranslated region. *Proc. Natl. Acad. Sci. USA*, 82: 2315-2319, 1985.
5. Kozak, M. Influences of mRNA secondary structure on initiation by eukaryotic ribosomes. *Proc. Natl. Acad. Sci. USA*, 85: 2850-2854, 1988.
6. Koromilas, A. E., Lazaris-Karatzas, A., and Sonenberg, N. mRNAs containing extensive secondary structure in their 5' non-coding region translate efficiently in cells overexpressing initiation factor eIF-4E. *EMBO J.*, 11: 4153-4158, 1992.
7. Rinker-Schaeffer, C. W., Graff, J. R., De Benedetti, A., Zimmer, S. G., and Rhoads, R. E. Decreasing the level of translation initiation factor 4E with antisense RNA causes reversal of ras-mediated transformation and tumorigenesis of cloned rat embryo fibroblasts. *Int. J. Cancer*, 55: 841-847, 1993.
8. Kevil, C. G., De Benedetti, A., Payne, D. K., Coe, L. L., Laroux, F. S., and Alexander, J. S. Translational regulation of vascular permeability factor by eukaryotic initiation factor 4E: Implications for tumor angiogenesis. *Int. J. Cancer*, 65: 785-790, 1996.
9. Nathan, C. A., Franklin, S., Abreo, F. W., Nassar, R., De Benedetti, A., and Glass, J. Analysis of surgical margins with the molecular marker eIF4E: a prognostic factor in patients with head and neck cancer. *J. Clin. Oncol.*, 17: 2909-2914, 1999.
10. Fukunaga, R., and Hunter, T. MNK1, a new MAP kinase-activated protein kinase, isolated by a novel expression screening method for identifying protein kinase substrates. *EMBO J.*, 16: 1921-1933, 1997.
11. Waskiewicz, A. J., Flynn, A., Proud, C. G., and Cooper, J. A. Mitogen-activated protein kinases activate the serine/threonine kinases Mnk1 and Mnk2. *EMBO J.*, 16: 1909-1920, 1997.
12. Wang, X., Flynn, A., Waskiewicz, A. J., Webb, B. L., Vrtes, R. G., Balnes, I. A., Cooper, J. A., and Proud, C. G. The phosphorylation of eukaryotic initiation factor eIF4E in response to phorbol esters, cell stresses, and cytokines is mediated by distinct MAP kinase pathways. *J. Biol. Chem.*, 273: 9373-9377, 1998.
13. Pyronnet, S., Imataka, H., Gingras, A. C., Fukunaga, R., Hunter, T., and Sonenberg, N. Human eukaryotic translation initiation factor 4G (eIF4G) recruits Mnk1 to phosphorylate eIF4E. *EMBO J.*, 18: 270-279, 1999.
14. Kleijn, M., Scheper, G. C., Voorma, H. O., and Thomas, A. A. Regulation of translation initiation factors by signal transduction. *Eur. J. Biochem.*, 253: 531-644, 1998.
15. Raught, B., and Gingras, A. C. eIF4E activity is regulated at multiple levels. *Int. J. Biochem. Cell Biol.*, 31: 43-57, 1999.
16. Takeuchi, K., Shibamoto, S., Nagamine, K., Shigemori, I., Omura, S., Kitamura, N., and Ito, F. Signaling pathways leading to transcription and translation cooperatively regulate the transient increase in expression of c-Fos protein. *J. Biol. Chem.*, 276: 26077-26083, 2001.
17. Kawasome, H., Papst, P., Webb, S., Keller, G. M., Johnson, G. L., Gelfand, E. W., and Terada, N. Targeted disruption of p70(s6k) defines its role in protein synthesis and rapamycin sensitivity. *Proc. Natl. Acad. Sci. USA*, 95: 5033-5038, 1998.
18. Christie, G. R., Hajdich, E., Hundal, H. S., Proud, C. G., and Taylor, P. M. Intracellular sensing of amino acids in *Xenopus laevis* oocytes stimulates p70 S6 kinase in a target of rapamycin-dependent manner. *J. Biol. Chem.*, 277: 9952-9957, 2002.
19. Hara, K., Yonezawa, K., Weng, Q. P., Kozlowski, M. T., Belham, C., and Avruch, J. Amino acid sufficiency and mTOR regulate p70 S6 kinase and eIF-4E BP1 through a common effector mechanism. *J. Biol. Chem.*, 273: 14484-14494, 1998.
20. Graves, L. M., Bornfeldt, K. E., Argast, G. M., Krebs, E. G., Kong, X., Lin, T. A., and Lawrence, J. C., Jr. cAMP- and rapamycin-sensitive regulation of the association of eukaryotic initiation factor 4E and the translational regulator PHAS-I in aortic smooth muscle cells. *Proc. Natl. Acad. Sci. USA*, 92: 7222-7226, 1995.
21. Merrick, W. C., and Hershey, J. W. B. The pathway and mechanism of eukaryotic protein synthesis. In: J. W. B. Hershey and M. B. Mathews (eds.), *Translational Control*, pp. 31-69. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory, 1998.
22. Kimball, S. R. Eukaryotic initiation factor eIF2. *Int. J. Biochem. Cell Biol.*, 31: 25-29, 1999.
23. Jagus, R., Joshi, B., and Barber, G. N. PKR, apoptosis and cancer. *Int. J. Biochem. Cell Biol.*, 31: 123-138, 1999.
24. Thomas, G., and Hall, M. N. TOR signalling and control of cell growth. *Curr. Opin. Cell Biol.*, 9: 782-787, 1997.
25. Gingras, A. C., Raught, B., and Sonenberg, N. Regulation of translation initiation by FRAP/mTOR. *Genes Dev.*, 15: 807-826, 2001.
26. Gingras, A. C., Gygi, S. P., Raught, B., Polakiewicz, R. D., Abraham, R. T., Hoekstra, M. F., Aebersold, R., and Sonenberg, N. Regulation of 4E-BP1 phosphorylation: a novel two-step mechanism. *Genes Dev.*, 13: 1422-1437, 1999.
27. Kumar, V., Pandey, P., Sabatini, D., Kumar, M., Majumder, P. K., Bhandi, A., Carmichael, G., Kufe, D., and Karbhandi, S. Functional interaction between RAFT1/FRAP/mTOR and protein kinase C $\delta$  in the regulation of cap-dependent initiation of translation. *EMBO J.*, 19: 1087-1097, 2000.
28. Yang, D. Q., and Kastan, M. B. Participation of ATM in insulin signaling through phosphorylation of eIF-4E-binding protein 1. *Nat. Cell Biol.*, 2: 893-899, 2000.
29. Liu, G., Zhang, Y., Bode, A. M., Ma, W. Y., and Dong, Z. Phosphorylation of 4E-BP1 is mediated by the p38/MSK1 pathway in response to UVB irradiation. *J. Biol. Chem.*, 277: 8810-8818, 2002.
30. Neshat, M. S., Meilinghoff, I. K., Tran, C., Stiles, B., Thomas, G., Petersen, R., Frost, P., Gibbons, J. J., Wu, H., and Sawyers, C. L. Enhanced sensitivity of PTEN-deficient tumors to inhibition of FRAP/mTOR. *Proc. Natl. Acad. Sci. USA*, 98: 10314-10319, 2001.
31. Selkovic, A., Hudson, C. C., Homma, J. L., Yin, P., Ottemess, D. M., Kamitz, L. M., and Abraham, R. T. A direct linkage between the phosphoinositide 3-kinase-AKT signaling pathway and the mammalian target of rapamycin in mitogen-stimulated and transformed cells. *Cancer Res.*, 60: 3504-3513, 2000.
32. Scott, P. H., and Lawrence, J. C., Jr. Attenuation of mammalian target of rapamycin activity by increased cAMP in 3T3-L1 adipocytes. *J. Biol. Chem.*, 273: 34496-34501, 1998.
33. Reynolds, I. T., Bodine, S. C., and Lawrence, J. C., Jr. Control of Ser2448 phosphorylation in the mammalian target of rapamycin by insulin and skeletal muscle load. *J. Biol. Chem.*, 277: 17657-17662, 2002.
34. Peterson, R. T., Beal, P. A., Comb, M. J., and Schreiber, S. L. FKBP12-rapamycin-associated protein (FRAP) autophosphorylates at serine 2481 under translationally repressive conditions. *J. Biol. Chem.*, 275: 7416-7423, 2000.
35. Peterson, R. T., Desai, B. N., Hardwick, J. S., and Schreiber, S. L. Protein phosphatase 2A interacts with the 70-kDa S6 kinase and is activated by inhibition of FKBP12-rapamycin-associated protein. *Proc. Natl. Acad. Sci. USA*, 96: 4438-4442, 1999.
36. McGrew, L. L., Dworkin-Rastl, E., Dworkin, M. B., and Richter, J. D. Poly(A) elongation during *Xenopus* oocyte maturation is required for translational recruitment and is mediated by a short sequence element. *Genes Dev.*, 3: 803-815, 1989.
37. Sheets, M. D., Wu, M., and Wickens, M. Polyadenylation of c-mos mRNA as a control point in *Xenopus* meiotic maturation. *Nature (Lond.)*, 374: 511-516, 1995.
38. Varnum, S. M., and Wormington, W. M. Deadenylation of maternal mRNAs during *Xenopus* oocyte maturation does not require specific cis-sequences: a default mechanism for translational control. *Genes Dev.*, 4: 2278-2286, 1990.
39. Gallie, D. R. The cap and poly(A) tail function synergistically to regulate mRNA translational efficiency. *Genes Dev.*, 5: 2108-2116, 1991.
40. Sachs, A. B., and Varani, G. Eukaryotic translation initiation: there are (at least) two sides to every story. *Nat. Struct. Biol.*, 7: 356-361, 2000.
41. Wolffe, A. P., and Meric, F. Coupling transcription to translation: a novel site for the regulation of eukaryotic gene expression. *Int. J. Biochem. Cell Biol.*, 28: 247-257, 1996.
42. Evdokimova, V. M., Wei, C. L., Sitkov, A. S., Simonenko, P. N., Lazarev, O. A., Vasilenko, K. S., Ustinov, V. A., Hershey, J. W., and Ovchinnikov, L. P. The major protein of messenger ribonucleoprotein particles in somatic cells is a member of the Y-box binding transcription factor family. *J. Biol. Chem.*, 270: 3186-3192, 1995.
43. Matsumoto, K., Meric, F., and Wolffe, A. P. Translational repression dependent on the interaction of the *Xenopus* Y-box protein FRGY2 with mRNA. Role of the cold shock domain, tail domain, and selective RNA sequence recognition. *J. Biol. Chem.*, 271: 22706-22712, 1996.



44. Evdokimova, V., Ruzanov, P., Imataka, H., Raught, B., Svitkin, Y., Orshinnikov, L. P., and Sonenberg, N. The major mRNA-associated protein YB-1 is a potent 5' cap-dependent mRNA stabilizer. *EMBO J.*, 20: 5491-5502, 2001.
45. Saito, H., Hayday, A. C., Wiman, K., Hayward, W. S., and Tonegawa, S. Activation of the *c-myc* gene by translocation: a model for translational control. *Proc. Natl. Acad. Sci. USA*, 80: 7476-7480, 1983.
46. Nanbru, C., Lafon, I., Audigier, S., Gensac, M. C., Vagner, S., Huez, G., and Prats, A. C. Alternative translation of the proto-oncogene *c-myc* by an internal ribosome entry site. *J. Biol. Chem.*, 272: 32061-32066, 1997.
47. Stoneley, M., Paulin, F. E., Le Quesne, J. P., Chappell, S. A., and Willis, A. E. *c-Myc* 5' untranslated region contains an internal ribosome entry segment. *Oncogene*, 16: 423-428, 1998.
48. Paulin, F. E., West, M. J., Sullivan, N. F., Whitney, R. L., Lyne, L., and Willis, A. E. Aberrant translational control of the *c-myc* gene in multiple myeloma. *Oncogene*, 13: 505-513, 1996.
49. Chappell, S. A., LeQuesne, J. P., Paulin, F. E., de Schootmeester, M. L., Stoneley, M., Soutar, R. L., Ralston, S. H., Helfrich, M. H., and Willis, A. E. A mutation in the *c-myc*-IRES leads to enhanced internal ribosome entry in multiple myeloma: a novel mechanism of oncogene de-regulation. *Oncogene*, 19: 4437-4440, 2000.
50. Signori, E., Bagni, C., Papa, S., Primerano, B., Rinaldi, M., Amaldi, F., and Fazio, V. M. A somatic mutation in the 5'UTR of *BRCA1* gene in sporadic breast cancer causes down-modulation of translation efficiency. *Oncogene*, 20: 4598-4600, 2001.
51. Liu, L., Dilworth, D., Gao, L., Morison, J., Summers, A., Lassam, N., and Hogg, D. Mutation of the *CDKN2A* 5' UTR creates an aberrant initiation codon and predisposes to melanoma. *Nat. Genet.*, 21: 128-132, 1999.
52. Savitsky, K., Platzer, M., Uziel, T., Gliad, S., Sartiel, A., Rosenthal, A., Eroy-Stein, O., Shilo, Y., and Rotman, G. Ataxia-telangiectasia: structural diversity of untranslated sequences suggests complex post-transcriptional regulation of *ATM* gene expression. *Nucleic Acids Res.*, 25: 1678-1684, 1997.
53. Brown, C. Y., Mize, G. J., Pineda, M., George, D. L., and Morris, D. R. Role of two upstream open reading frames in the translational control of oncogene *mdm2*. *Oncogene*, 18: 5631-5637, 1999.
54. Sobczak, K., and Krzyzosiak, W. J. Structural determinants of *BRCA1* translational regulation. *J. Biol. Chem.*, 277: 17349-17358, 2002.
55. Arrick, B. A., Grendell, R. L., and Griffin, L. A. Enhanced translational efficiency of a novel transforming growth factor  $\beta$  mRNA in human breast cancer cells. *Mol. Cell. Biol.*, 14: 619-628, 1994.
56. Myers, S. M., Eng, C., Ponder, B. A., and Mulligan, L. M. Characterization of *RET* proto-oncogene 3' splicing variants and polyadenylation sites: a novel C-terminus for *RET*. *Oncogene*, 11: 2039-2045, 1995.
57. Byrne, J. A., Tomasello, C., Rouyer, N., Bellocq, J. P., Rio, M. C., and Basset, P. The tissue inhibitor of metalloproteinases-3 gene in breast carcinoma: identification of multiple polyadenylation sites and a stromal pattern of expression. *Mol. Med.*, 1: 418-427, 1995.
58. Moscow, J. A., He, R., Gudas, J. M., and Cowan, K. H. Utilization of multiple polyadenylation signals in the human *RHOA* protooncogene. *Gene (Amst.)*, 144: 229-236, 1994.
59. Senterre-Lesenfants, S., Alag, A. S., and Sobel, M. E. Multiple mRNA species are generated by alternate polyadenylation from the human *calmodulin-I* gene. *J. Cell. Biochem.*, 58: 445-454, 1995.
60. Lazaris-Karatzas, A., Montine, K. S., and Sonenberg, N. Malignant transformation by a eukaryotic initiation factor subunit that binds to mRNA 5' cap. *Nature (Lond.)*, 345: 544-547, 1990.
61. De Benedetti, A., and Rhoads, R. E. Overexpression of eukaryotic protein synthesis initiation factor 4E in HeLa cells results in aberrant growth and morphology. *Proc. Natl. Acad. Sci. USA*, 87: 8212-8216, 1990.
62. Polunovsky, V. A., Rosenwald, I. B., Tan, A. T., White, J., Chiang, L., Sonenberg, N., and Blitman, P. B. Translational control of programmed cell death: eukaryotic translation initiation factor 4E blocks apoptosis in growth-factor-restricted fibroblasts with physiologically expressed or de-regulated *Myc*. *Mol. Cell. Biol.*, 16: 6573-6581, 1996.
63. Miyagi, Y., Sugiyama, A., Asai, A., Okazaki, T., Kuchino, Y., and Kerr, S. J. Elevated levels of eukaryotic translation initiation factor eIF-4E mRNA in a broad spectrum of transformed cell lines. *Cancer Lett.*, 97: 247-252, 1995.
64. Kerekatta, V., Smiley, K., Hu, B., Smith, A., Gelder, F., and De Benedetti, A. The proto-oncogene/translation factor eIF4E: a survey of its expression in breast carcinomas. *Int. J. Cancer*, 64: 27-31, 1995.
65. Li, B. D., Liu, L., Dawson, M., and De Benedetti, A. Overexpression of eukaryotic initiation factor 4E (eIF4E) in breast carcinoma. *Cancer (Phila.)*, 79: 2385-2390, 1997.
66. Sorrells, D. L., Black, D. R., Meschonat, C., Rhoads, R., De Benedetti, A., Gao, M., Williams, B. J., and Li, B. D. Detection of eIF4E gene amplification in breast cancer by competitive PCR. *Ann. Surg. Oncol.*, 5: 232-237, 1998.
67. Li, B. D., McDonald, J. C., Nassar, R., and De Benedetti, A. Clinical outcome in stage I to III breast carcinoma and eIF4E overexpression. *Ann. Surg.*, 227: 758-761; discussion, 761-763, 1998.
68. Wang, S., Rosenwald, I. B., Hutzler, M. J., Pihan, G. A., Savas, L., Chen, J. J., and Woda, B. A. Expression of the eukaryotic translation initiation factors 4E and 2a in non-Hodgkin's lymphomas. *Am. J. Pathol.*, 155: 247-255, 1999.
69. Fukuchi-Shimogori, T., Ishii, I., Kashiwagi, K., Mashiba, H., Ekimoto, H., and Igarashi, K. Malignant transformation by overproduction of translation initiation factor eIF4G. *Cancer Res.*, 57: 5041-5044, 1997.
70. Eberle, J., Krasagakis, K., and Orfanos, C. E. Translation initiation factor eIF-4A1 mRNA is consistently overexpressed in human melanoma cells *in vitro*. *Int. J. Cancer*, 71: 398-401, 1997.
71. Shuda, M., Kondoh, N., Tanaka, K., Ryo, A., Wakatsuki, T., Hada, A., Goseki, N., Igarashi, T., Hattusa, K., Aihara, T., Horuchi, S., Shichita, M., Yamamoto, N., and Yamamoto, M. Enhanced expression of translation factor mRNAs in hepatocellular carcinoma. *Anticancer Res.*, 20: 2489-2494, 2000.
72. Nupponen, N. N., Pirkka, K., Kakkola, L., Tanner, M., Persson, K., Borg, A., Isola, J., and Visakorpi, T. Amplification and overexpression of p40 subunit of eukaryotic translation initiation factor 3 in breast and prostate cancer. *Am. J. Pathol.*, 154: 1777-1783, 1999.
73. Rothe, M., Ko, Y., Albers, P., and Wernert, N. Eukaryotic initiation factor 3 p110 mRNA is overexpressed in testicular seminomas. *Am. J. Pathol.*, 157: 1597-1604, 2000.
74. Barlund, M., Forozan, F., Kononen, J., Bubendorf, L., Chen, Y., Bittner, M. L., Torhorst, J., Haas, P., Bucher, C., Sauter, G., Kallioniemi, O. P., and Kallioniemi, A. Detecting activation of ribosomal protein S6 kinase by complementary DNA and tissue microarray analysis. *J. Natl. Cancer Inst.* (Bethesda), 92: 1252-1259, 2000.
75. Topalian, S. L., Kaneko, S., Gorzales, M. I., Bond, G. L., Ward, Y., and Manley, J. L. Identification and functional characterization of neo-poly(A) polymerase, an RNA processing enzyme overexpressed in human tumors. *Mol. Cell. Biol.*, 21: 5614-5623, 2001.
76. Scorilas, A., Toller, M., Ardavanis, A., Courtis, N., Dimitriadis, E., Yotis, J., Tsiapalis, C. M., and Trangas, T. Polyadenylate polymerase enzymatic activity in mammary tumor cytosols: a new independent prognostic marker in primary breast cancer. *Cancer Res.*, 60: 5427-5433, 2000.
77. Janz, M., Harbeck, N., Deltmar, P., Berger, U., Schmidt, A., Jurchott, K., Schmitt, M., and Royer, H. D. Y-box factor YB-1 predicts drug resistance and patient outcome in breast cancer independent of clinically relevant tumor biologic factors HER2, uPA and PAI-1. *Int. J. Cancer*, 97: 278-282, 2002.
78. Shibahara, K., Sugio, K., Osaki, T., Uchiyama, T., Maehara, Y., Kohno, K., Yasumoto, K., Sugimachi, K., and Kuwano, M. Nuclear expression of the Y-box binding protein, YB-1, as a novel marker of disease progression in non-small cell lung cancer. *Clin. Cancer Res.*, 7: 3151-3155, 2001.
79. Kamura, T., Yahata, H., Amada, S., Ogawa, S., Sonoda, T., Kobayashi, H., Mitsumoto, M., Kohno, K., Kuwano, M., and Nakano, H. Is nuclear expression of Y box-binding protein-1 a new prognostic factor in ovarian serous adenocarcinoma? *Cancer (Phila.)*, 85: 2450-2454, 1999.
80. Bargou, R. C., Jurchott, K., Wagener, C., Bergmann, S., Metzner, S., Bommer, K., Mapara, M. Y., Winzer, K. J., Dietel, M., Dorkan, B., and Royer, H. D. Nuclear localization and increased levels of transcription factor YB-1 in primary human breast cancers are associated with intrinsic *MDR1* gene expression. *Nat. Med.*, 3: 447-450, 1997.
81. Aoki, M., Blazek, E., and Vogt, P. K. A role of the kinase mTOR in cellular transformation induced by the oncoproteins P3k and Akt. *Proc. Natl. Acad. Sci. USA*, 98: 136-141, 2001.
82. Child, S. J., Miller, M. K., and Geballe, A. P. Cell type-dependent and -independent control of HER-2/neu translation. *Int. J. Biochem. Cell Biol.*, 31: 201-213, 1999.

83. Jefferies, H. B., Reinhard, C., Kozma, S. C., and Thomas, G. Rapamycin selectively represses translation of the "polypyrimidine tract" mRNA family. *Proc. Natl. Acad. Sci. USA*, 91: 4441-4445, 1994.
84. Terada, N., Patel, H. R., Takase, K., Kohno, K., Naim, A. C., and Gelfand, E. W. Rapamycin selectively inhibits translation of mRNAs encoding elongation factors and ribosomal proteins. *Proc. Natl. Acad. Sci. USA*, 91: 11477-11481, 1994.
85. Jefferies, H. B., Fumagalli, S., Dennis, P. B., Reinhard, C., Pearson, R. B., and Thomas, G. Rapamycin suppresses 5'TOP mRNA translation through inhibition of p70S6k. *EMBO J.*, 16: 3693-3704, 1997.
86. Beretta, L., Gingras, A. C., Svitkin, Y. V., Hall, M. N., and Sonenberg, N. Rapamycin blocks the phosphorylation of 4E-BP1 and inhibits cap-dependent initiation of translation. *EMBO J.*, 15: 658-664, 1996.
87. Hidalgo, M., and Rowinsky, E. K. The rapamycin-sensitive signal transduction pathway as a target for cancer therapy. *Oncogene*, 19: 6880-6888, 2000.
88. Hosoi, H., Dilling, M. B., Shikata, T., Liu, L. N., Shu, L., Ashmun, R. A., Germain, G. S., Abraham, R. T., and Houghton, P. J. Rapamycin causes poorly reversible inhibition of mTOR and induces p53-independent apoptosis in human rhabdomyosarcoma cells. *Cancer Res.*, 59: 888-894, 1999.
89. Huang, S., and Houghton, P. J. Resistance to rapamycin: a novel anticancer drug. *Cancer Metastasis Rev.*, 20: 69-78, 2001.
90. Georger, B., Kerr, K., Tang, C. B., Fung, K. M., Powell, B., Sutton, L. N., Phillips, P. C., and Janss, A. J. Antitumor activity of the rapamycin analog CCI-779 in human primitive neuroectodermal tumor/medulloblastoma models as single agent and in combination chemotherapy. *Cancer Res.*, 61: 1527-1532, 2001.
91. Gibbons, J. J., Discafani, C., Peterson, R., Hernandez, R., Skotnick, J., and Frost, P. The effect of CCI-779, a novel macrolide anti-tumor agent, on the growth of human tumor cells *in vitro* and in nude mouse xenografts *in vivo*. *Proc. Am. Assoc. Cancer Res.*, 40: 301, 1999.
92. Hidalgo, M., Rowinsky, E., Erlichman, C., Marshall, B., Marks, R., Edwards, T., and Buckner, J. J. A Phase I and pharmacological study of CCI-779 cycle inhibitor. *Ann. Oncol.*, 11 (Suppl. 4): 133, 2001.
93. Alexandre, J., Raymond, E., Depenbrock, H., Mekhaldi, S., Angevin, E., Pallet, C., Hanauske, A., Frisch, J., Feussner, A., and Armand, J. P. CCI-779, a new rapamycin analog, has antitumor activity at doses inducing only mild cutaneous effects and mucositis: early results of an ongoing Phase I study. *Proceedings of the 1999 AACR-NCI-EORTC International Conference, Clin. Cancer Res.*, 5 (Suppl.): 3730s, 1999.
94. Chan, S., Johnston, S., Scheulen, M. E., Mross, K., Morant, A., Lehr, A., Feussner, A., Berger, M., and Kirsch, T. First report: a Phase 2 study of the safety and activity of CCI-779 for patients with locally advanced or metastatic breast cancer failing prior chemotherapy. *Proc. Am. Soc. Clin. Oncol.*, 21: 44a, 2002.
95. Atkins, M. B., Hidalgo, M., Stadler, W., Logan, T., Dutcher, J. P., Hudes, G., Park, Y., Marshall, B., Boni, J., and Dukat, G. A randomized double-blind Phase 2 study of intravenous CCI-779 administered weekly to patients with advanced renal cell carcinoma. *Proc. Am. Soc. Clin. Oncol.*, 21: 10a, 2002.
96. Smith, S. G., Trinh, C. M., Inge, L. J., Thomas, G., Cloughesy, T. F., Sawyers, C. L., and Mischel, P. S. PTEN expression status predicts glioblastoma cell sensitivity to CCI-779. *Proc. Am. Assoc. Cancer Res.*, 43: 335, 2002.
97. Yu, K., Toral-Barza, L., Discafani, C., Zhang, W. G., Skotnick, J., Frost, P., and Gibbons, J. J. mTOR, a novel target in breast cancer: the effect of CCI-779, an mTOR inhibitor, in preclinical models of breast cancer. *Endocr. Relat. Cancer*, 8: 249-258, 2001.
98. Dilling, M. B., Germain, G. S., Dudkin, L., Jayaraman, A. L., Zhang, X., Harwood, F. C., and Houghton, P. J. 4E-binding proteins, the suppressors of eukaryotic initiation factor 4E, are downregulated in cells with acquired or intrinsic resistance to rapamycin. *J. Biol. Chem.*, 277: 13907-13917, 2002.
99. Guba, M., von Breitenbuch, P., Steinbauer, M., Koehl, G., Fiegel, S., Hornung, M., Bruns, C. J., Zuelke, C., Farkas, S., Anthuber, M., Jauch, K. W., and Gelsner, E. K. Rapamycin inhibits primary and metastatic tumor growth by antiangiogenesis: Involvement of vascular endothelial growth factor. *Nat. Med.*, 8: 128-135, 2002.
100. Lane, H. A., Schell, C., Thauer, A., O'Reilly, T., and Wood, J. Anti-angiogenic activity of RAD001, an orally active anticancer agent. *Proc. Am. Assoc. Cancer Res.*, 43: 184, 2002.
101. Mashima, Y., Sudhakar, A., Lively, J. C., Ueki, K., Kharbanda, S., Kahn, C. R., Sonenberg, N., Hynes, R. O., and Kalluri, R. Tumstatin, an endothelial cell-specific inhibitor of protein synthesis. *Science (Wash. DC)*, 295: 140-143, 2002.
102. Caygill, C. P., Charlett, A., and Hill, M. J. Fat, fish, fish oil and cancer. *Br. J. Cancer*, 74: 159-164, 1996.
103. Falconer, J. S., Ross, J. A., Fearon, K. C., Hawkins, R. A., O'Riordan, M. G., and Carter, D. C. Effect of eicosapentaenoic acid and other fatty acids on the growth *in vitro* of human pancreatic cancer cell lines. *Br. J. Cancer*, 69: 828-832, 1994.
104. Noguchi, M., Mhami, M., Yegasaki, R., Knosha, K., Earashi, M., Ktagawa, H., Taniya, T., and Miyazaki, I. Chemoprevention of DMBA-induced mammary carcinogenesis in rats by low-dose EPA and DHA. *Br. J. Cancer*, 75: 348-353, 1997.
105. Palakurthi, S. S., Fluckiger, R., Aktas, H., Changolkar, A. K., Shahsafaei, A., Harnett, S., Kilic, E., and Halperin, J. A. Inhibition of translation initiation mediates the anticancer effect of the n-3 polyunsaturated fatty acid eicosapentaenoic acid. *Cancer Res.*, 60: 2919-2925, 2000.
106. Aktas, H., Fluckiger, R., Acosta, J. A., Savage, J. M., Palakurthi, S. S., and Halperin, J. A. Depletion of intracellular  $Ca^{2+}$  stores, phosphorylation of eIF2 $\alpha$ , and sustained inhibition of translation initiation mediate the anticancer effects of clotrimazole. *Proc. Natl. Acad. Sci. USA*, 95: 8280-8285, 1998.
107. Mhashilkar, A. M., Schrock, R. D., Hindi, M., Liao, J., Sieger, K., Kourouma, F., Zou-Yang, X. H., Onishi, E., Takh, O., Vedvick, T. S., Fanger, G., Stewart, L., Watson, G. J., Snary, D., Fisher, P. B., Seek, T., Roth, J. A., Ramesh, R., and Chada, S. Melanoma differentiation associated gene-7 (mda-7): a novel anti-tumor gene for cancer gene therapy. *Mol. Med.*, 7: 271-282, 2001.
108. Su, Z. Z., Madireddi, M. T., Lin, J. J., Young, C. S., Kitada, S., Reed, J. C., Goldstein, N. I., and Fisher, P. B. The cancer growth suppressor gene mda-7 selectively induces apoptosis in human breast cancer cells and inhibits tumor growth in nude mice. *Proc. Natl. Acad. Sci. USA*, 95: 14400-14405, 1998.
109. Seek, T., Mhashilkar, A., Chada, S., Branch, C., Roth, J. A., and Ramesh, R. Tumor-suppressive effects by adenovirus-mediated mda-7 gene transfer in non-small cell lung cancer cell *in vitro*. *Gene Ther.*, 7: 2051-2057, 2000.
110. Pataer, A., Vorburger, S. A., Barber, G. N., Chada, S., Mhashilkar, A. M., Zou-Yang, H., Stewart, A. L., Balachandran, S., Roth, J. A., Hunt, K. K., and Swisher, S. G. Adenoviral transfer of the melanoma differentiation-associated gene 7 (mda7) induces apoptosis of lung cancer cells via up-regulation of the double-stranded RNA-dependent protein kinase (PKR). *Cancer Res.*, 62: 2239-2243, 2002.
111. Ito, T., Wamken, S. P., and May, W. S. Protein synthesis inhibition by flavonoids: roles of eukaryotic initiation factor 2 $\alpha$  kinases. *Biochem. Biophys. Res. Commun.*, 265: 589-594, 1999.
112. Eberle, J., Fecker, L. F., Bittner, J. U., Orfanos, C. E., and Geilen, C. C. Decreased proliferation of human melanoma cell lines caused by antisense RNA against translation factor eIF-4A1. *Br. J. Cancer*, 86: 1957-1962, 2002.
113. Polunovsky, V. A., Gingras, A. C., Sonenberg, N., Peterson, M., Tan, A., Rubins, J. B., Manivel, J. C., and Bittman, P. B. Translational control of the antiapoptotic function of Ras. *J. Biol. Chem.*, 275: 24776-24780, 2000.
114. D'Cunha, J., Kratzke, M. G., Alter, M. D., Polunovsky, V. A., Bittman, P. B., and Kratzke, R. A. Over-expression of the translational repressor 4E-BP1 inhibits NSCLC tumorigenicity *in vivo*. *Proc. Am. Assoc. Cancer Res.*, 43: 816-817, 2002.
115. DeFatta, R. J., Nathan, C. A., and De Benedetti, A. Antisense RNA to eIF4E suppresses oncogenic properties of a head and neck squamous cell carcinoma cell line. *Laryngoscope*, 110: 928-933, 2000.
116. Herbert, T. P., Fahraeus, R., Prescott, A., Lane, D. P., and Proud, C. G. Rapid induction of apoptosis mediated by peptides that bind initiation factor eIF4E. *Curr. Biol.*, 10: 793-798, 2000.
117. DeFatta, R. J., Li, Y., and De Benedetti, A. Selective killing of cancer cells based on translational control of a suicide gene. *Cancer Gene Ther.*, 9: 573-578, 2002.

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☒ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☒ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☒ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**